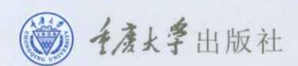


LATENT VARIABLE MODELING USING M*plus*

潜变量建模与Mplus应用·进阶篇



王孟成 毕向阳 著



图书在版编目（CIP）数据

潜变量建模与M*plus* 应用·进阶篇 / 王孟成，毕向阳著．—重庆：重庆大学出版社，2018.4

（万卷方法）

ISBN 978-7-5689-0816-0

Ⅰ．①潜…　Ⅱ．①王…②毕…　Ⅲ．①统计模型—统计分析　Ⅳ.①C81

中国版本图书馆CIP数据核字（2017）第245490号

潜变量建模与M*plus* 应用·进阶篇

王孟成　毕向阳　著

策划编辑：林佳木

责任编辑：林佳木　　版式设计：林佳木

责任校对：邬小梅　　责任印制：赵　晟

*

重庆大学出版社出版发行

出版人：易树平

社址：重庆市沙坪坝区大学城西路21号

邮编：401331

电话：（023）88617190　88617185（中小学）

传真：（023）88617186　88617166

网址：http：//www.cqup.com.cn

邮箱：fxk@cqup.com.cn（营销中心）

全国新华书店经销

重庆升光电力印务有限公司印刷

*

开本：787mm×1092mm　1／16　印张：21.5　字数：462千　插页：16开6页

2018年4月第1版　　2018年4月第1次印刷

印数：1—4000

ISBN 978-7-5689-0816-0

前言

《潜变量建模与M*plus* 应用·进阶篇》是2014年出版的《潜变量模型与M*plus* 应用·基础篇》的姊妹书，主要涉及更复杂的潜变量模型。全书共10章，分为混合模型篇（第1至5章）、多水平模型篇（第6至9章）和贝叶斯结构方程模型篇（第10章）。

混合模型（Mixture Modeling）是最近几年开始流行的潜变量建模方法，其目的是在建模过程中考虑群体的异质性，让传统建模方法更加精确。多水平模型对于很多学科来说非常重要，在管理学、教育学等领域流行多年，尽管已存在多种处理多水平模型的软件，M*plus* 却是其中最灵活的一个，不仅可以将文中提及的多种潜变量模型置于多水平的框架之下，而且也能够将若干前沿的建模和检验方法方便地实现。最后一章则用于介绍潜变量建模的贝叶斯方法。

和基础篇一样，本书的定位是应用取向的。在介绍模型时保留了基本的、必要的数学公式，但我们同时尝试使用文字解释，让惧怕数学公式的读者也能理解模型的基本思想。每种模型的介绍都配备实际分析的数据和M*plus* 语句，帮助读者理解建模过程、软件使用和结果解读。

据我们所知，本书是国内第一本介绍混合模型的教材，也是第一本系统介绍M*plus* 分析多水平模型的教材。由于种种原因，书中难免存在一些不足和错误，希望各位读者给予反馈，以期将来再版或重印时补足或纠正。

随着潜变量建模方法的普及，国内M*plus* 使用者正在迅速扩大。为“基础篇”读者交流创建的QQ群（219343650）成员已近1600人，希

望“进阶篇”的出版能进一步提升大家的研究质量，推进国内有关潜变量建模的应用。

“进阶篇”的读者交流QQ群（453434352）。

王孟成　毕向阳

目　录

第一部分　混合模型

1　潜类别模型

随着方法学的发展，以变量为中心和以个体为中心的方法之间的界限已开始模糊，有逐渐融合的趋势。例如，因子混合模型（Factor Mixture Modeling，FMM）将类别潜变量和连续潜变量融合到一个分析模型中，这也是目前潜变量建模界的趋势。

《潜变量建模与M*plus* 应用·基础篇》介绍的方法基本上都在因子分析的框架内，即潜变量为连续型变量的模型。同时，分析的对象是指标或变量，因此也称作变量中心化的方法（Variable-Centered）。本书的前面几章将关注以被试或个体为中心（Person-Centered）的分析方法。

1.1 潜类别模型概述

潜在类别模型（Latent Class Model，LCM；Lazarsfeld & Henry，1968）或潜在类别分析（Latent Class Analysis，LCA）是通过间断的潜变量即潜在类别（Class）变量来解释外显指标间的关联，使外显指标间的关联通过潜在类别变量来解释，进而维持其局部独立性的统计方法（见图1-1）。其基本假设是，外显变量各种反应的概率分布可以由少数互斥的潜在类别变量来解释，每种类别对各外显变量的反应选择都有特定的倾向（邱皓政，2008；Collins & Lanza，2010）。与潜在类别分析非常相似的是潜在剖面分析（Latent Profile Analysis，LPA），区别在于前者处理分类观测变量，后者分析连续观测变量。

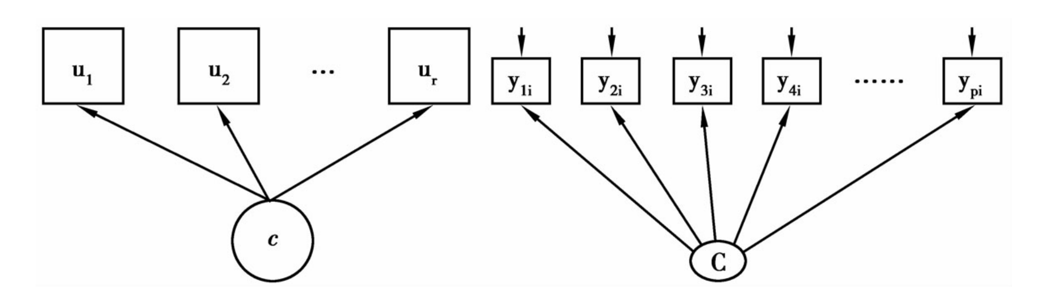


图1-1 LCM和LPM示意图

近年来LCM在心理学、预防医学、精神病学、市场营销、组织管理等诸多领域广为使用（e．g．，Carragher et al，2009；Lanza＆Rhoades，2011；Wang＆Hanges，2011；张洁婷，焦璨，张敏强，2010）。例如，Carragher et al（2009）在12180个全美代表性样本中将抑郁症状划分成四个类别：严重抑郁组（Severely Depressed，40.9％），躯体症状组（Psychosomatic，30.6％），认知情感组（Cognitive-Emotional，10.2％）和健康组（Non-depressed，18.3％）。

1.2 列联表及其概率分布

列联表（Contingency）是最常见的用于描述类别变量间关系的形式。表1-1是一个常见的2×2列联表，如果存在多个二分变量，任意两个变量便可组成一个2×2列联表，多个变量之间可以组成多重列联表。列联表涉及三种类型的概率分布：联合概率（Joint Probability），边际概率（Marginal Probability）和条件概率（Conditional Probability）。

表1-1 2×2列联表

			条目 A		
			是=1	否=0	Total
条目 B	是=1	计数 联合概率 条件概率(列) 条件概率(行)	$n_{A=1,B=1}=306$ $P(A=1\cap B=1)=.306$ $P(B=1\mid A=1)=.516$ $P(A=1\mid B=1)=.594$	$n_{A=0,B=1}=209$ $P(A=0\cap B=1)=.209$ $P(B=1\mid A=0)=.514$ $P(A=0\mid B=1)=.406$	$n_{B=1}=515$ $P(B=1)=.515$
	否=0	计数 联合概率 条件概率(列) 条件概率(行)	$n_{A=1,B=0}=287$ $P(A=1\cap B=0)=.287$ $P(B=0\mid A=1)=.484$ $P(A=1\mid B=0)=.592$	$n_{A=0,B=0}=198$ $P(A=0\cap B=0)=.198$ $P(B=0\mid A=0)=.486$ $P(A=0\mid B=0)=.408$	$N_{B=0}=485$ $P(B=0)=.485$
	Total	边际总和 边际概率	$n_{A=1}=593$ $P(A=1)=.593$	$n_{A=0}=407$ $P(A=0)=.407$	$n_{AB}=1000$ $P(A\cap B)=1$

注：基于A 事件发生B 事件发生的概率：

$P(B=1\mid A=1)=P(A=1\cap B=1)/P(A=1)=n_{A=1,\ B=1}/n_{A=1}=306/593$ $=.516$

$P(B=0\mid A=1)=P(A=1\cap B=0)/P(A=1)=n_{A=1,\ B=0}/n_{A=1}=287/593$ $=.484$

$P(B=1\mid A=0)=P(A=0\cap B=1)/P(A=0)=n_{A=0,\ B=1}/n_{A=0}=209/515$ $=.514$

$P(B=0\mid A=0)=P(A=0\cap B=0)/P(A=0)=n_{A=0,\ B=0}/n_{A=0}=198/407$ $=.486$

基于B 事件发生A 事件发生的概率：

$P(A=1\mid B=1)=P(A=1\cap B=1)/P(B=1)=n_{A=1,\ B=1}/n_{B=1}=306/515$ $=.594$

$P(A=1\mid B=0)=P(A=1\cap B=0)/P(B=0)=n_{A=1,\ B=0}/n_{B=0}=287/485$ $=.592$

$P(A=0 | B=1) = P(A=0 \cap B=1) / P(B=1) = n_{A=0,\ B=1} / n_{B=1} = 209 / 515$

$= .406$

$P(A=0 | B=0) = P(A=0 \cap B=0) / P(B=0) = n_{A=0,\ B=0} / n_{B=0} = 198 / 485$

$= .408$

联合概率指两个以上事件同时发生的概率，记作$P(A \cap B)$。表中的每个单元格的计数和概率即表示2种不同事件同时发生的情况，联合概率之和等于1。行或列的**边际概率** 是行或列联合概率之和。**条件概率** 指基于事件A发生，B事件发生的概率，记作$P(B / A) = P(A \cap B) / P(A) = n_{AB} / n_A$。通俗地说，条件概率就是属于某个变量的某个特定类别的个体在另一个变量各类别上的反应概率。例如，在条目A上选“是＝1”的个体在条目B的两个选项1和0上的分布情况；或者在条目B上选“否＝0”的个体在条目A的两个选项1和0上的分布情况。

如果A的条件分布概率在B的各个水平上相等，说明这两个变量间不存在统计关联即统计独立，因此可以通过条件概率来描述两个变量的关系。两个变量间统计独立等价于联合概率等于边际概率之积，$P(A=i \cap B=j) = P(A=i) \times P(B=j)$。

列联表统计独立性检验最常用的是皮尔逊卡方（Pearson Chi-Square）和似然比（Likelihood-Ratio）检验：

$$\chi^2 = \sum \frac{(n_{ij} - \mu_{ij})^2}{\mu_{ij}} \quad (1.1)$$

$$G^2 = 2 \sum n_{ij} log\left(\frac{n_{ij}}{u_{ij}}\right) \quad (1.2)$$

n_{ij} 为单元格观测频数，μ_{ij} 表示单元格理论频数，

$$\mu_{ij} = n\, p_{i+}\, p_{+j} = n\left(\frac{n_{i+}}{n}\right)\left(\frac{n_{+j}}{n}\right) = \frac{n_{i+}n_{+j}}{n} \qquad (1.3)$$

p_{i+} 和p_{+j} 分别为行和列的边际概率，n_{i+} 和n_{+j} 为行和列的边际总和。当样本量足够大时，G^2 服从卡方分布。

1.3　潜类别分析原理

LCM是根据个体在观测指标上的反应模式即不同的联合概率来进行参数估计的统计方法。例如，一份数学测验有10道判断题，数学能力高的个体能全部正确地回答所有题目，能力低的学生只能正确回答容易的题目，能力中等的学生能正确回答全部容易的和部分困难的题目。不同能力水平的学生在正确回答不同难易水平的题目时表现出某种相似性，因此通过学生回答题目的情况可以将其分为不同的能力水平组。LCM分析的逻辑就是根据个体在观测项目上的反应模式将其分类。

1.3.1　数学表达式

（1）潜类别分析模型

可以从方差分析的角度理解LCM。方差分析的特点是将方差分解成不同的来源，常见的有组间vs. 组内和被试间vs. 被试内。在LCM中，

可以将方差分解为类别内和类别间（Sterba，2013）。

根据局部独立性（local independence）假设，类别内的任意两个观测指标间的关联已通过潜类别变量解释，所以它们之间已没有关联。根据独立事件联合发生的概率等于单独发生概率之积的原理，在每个类别内部，多个两点计分项目的联合概率可以表示为：

$$p(y_i \mid c_i = k) = \prod_{j=1}^{J} p(y_{ij} \mid c_i = k) \qquad (1.4)$$

上式中，y_i 表示个体i 在指标j的两个选项y ＝1或y ＝0的得分。下标j 表示2点计分的指标，c 为潜类别变量，有k 个水平。

同时考虑多个类别水平时，上式扩展为：

$$p(y_i) = \sum_{k=1}^{K} p(c_i = k) p(y_i \mid c_i = k) \qquad (1.5)$$

p（c_i ＝k） 表示某一类别组k 所占总体的比率，亦称潜类别概率。

（2）潜剖面分析模型

当观测指标为连续型变量时，LCM称作潜在剖面分析（Latent Profile Modeling，LPM）。此时，将连续指标的方差分解为类别／剖面间和类别／剖面内方差（Lazarsfeld＆Henry，1968）：

$$\sigma_i^2 = \sum_{k=1}^{K} P(c_i = k)(\mu_{ik} - \mu_i)^2 + \sum_{k=1}^{K} P(c_i = k)\sigma_{ik}^2 \qquad (1.6)$$

其中μ_{ik}和σ_{ik}^2为剖面k内指标i的均值和方差。$P(c_i=k)$为类别概率，即每个类别个体占全体的比例。当满足局部独立性和同质假设时，上式简化为：

$$f(y_i)=\sum_{k=1}^{K}P(c_i=k)f(y_i\mid c_i=k) \tag{1.7}$$

1.3.2　类别概率和条件概率

在LCM中，有两个非常重要的参数：潜类别概率和条件概率。潜在类别概率类似于FA中的解释方差比例。在FA中，解释方差比例说明每个因子在解释数据时所占的分量。LCM的潜在类别概率则用于将样本分成不同比例的类别。换句话说，潜在类别概率就是用于说明各个类别的人数比例。例如，根据1000个被试在4个观测指标上的得分情况将其分成3个潜在类别，潜在类别概率为70％，25％和5％，表示第1类有700个被试，占70％；第2类有250个被试，占25％；第3类有50个被试，占5％。

条件概率指潜类别组内的个体在外显指标上的作答概率。例如，类别潜变量C有3个类别C1，C2和C3，外显指标A有3个选项，C1的条件概率就是计算C1内的个体在A的三个选项上的选择比例。以此类推，C2和C3也可以计算各自在A上的条件概率，共9个（每个类别有三个）。由于T1内的个体必然在A的三个不同选项的某个特定选项上选择，所以类别内的条件概率总和等于1，公式如下。

$$p(c_i=k)=\sum(p(c_i=1)+p(c_i=2)+\cdots+p(c_i=k))=1 \tag{1.8}$$

表1-2中呈现了下文探索性LCA分析时4个类别在前5个指标选项1上的条件概率和潜类别概率。

表1-2　条件概率和潜类别概率

潜类别	条件概率					潜类别概率
	T1	T2	T3	T4	T5	
C1	.390	.435	.276	.184	.182	.2054
C2	.903	.963	.951	.872	.943	.2946
C3	.493	.710	.468	.339	.538	.3357
C4	.739	.790	.883	.551	.883	.1643

注：T1～T5为五个项目；由于每个项目只有两个选项，所以表中只呈现了在一个选项上的条件概率，在第二个选项上的条件概率为1减表中数值。

条件概率与因子负荷类似，表达潜变量与外显变量之间关系的强弱。各潜在类别的概率总和及每个外显变量的条件概率总和都为1，这是LCA模型的基本限制条件。也可以根据特定研究的相关理论对参数进行限定。

一旦最优的LCA模型拟合成功，就需要将每个个体归入不同的潜在类别。换句话说，就是确定每个个体的潜类别属性（Class Membership）。在LCA中，采用的分类依据是贝叶斯后验概率（Posterior Probability），公式如下。

$$p(c_i = k \mid y_i) = \frac{p(c_i = k) f(y_i \mid c_i = k)}{f(y_i)} \qquad (1.9)$$

后验类别概率是根据个体的作答类型，在LCA拟合后估计得来，其值表示个体属于某一类别的概率。常见的分类方法有3种：莫代尔分配法（Modal assignment）、比例分配法（Proportional assignment）和虚拟类别法。莫代尔分配法根据个体后验概率的最大值将其归入特

定类别。例如，某个体A在4个类别上的后验概率分别为0.80，0.10，0.05和0.05。根据此值，A在第一类别中的概率值最高，所以应将其归入第一类。比例分配法则不直接分类而使用后验概率作为权重。虚拟类别法从后验概率分布中随机抽取一定次数的值，然后采用类似多重插补的方法得到平均的结果。M*plus* 采用的是莫代尔分配法，后验概率和个体类别归属变量可通过表1-3中的命令获得：

表1-3　LCM分类概率的M*plus* 语句

```
SAVEDATA：FILE＝ptsd-lca-4. txt；（保存文件名）。
    SAVE＝cprob；（要求报告后验概率）。
```

1.3.3　同质性和潜类别距离

同质性（Homogeneity）是指特定潜类别内部个体在外显变量上反应类型的一致性。同质性高说明同一类别内个体有着相同的反应类型，反之，说明反应类型差异较大。

潜类别间隔（Latent Class Separation）是指潜类别间差异的大小，反映在项目反应概率上则表现为不同类别个体间在所有观测指标上存在显著的差别（Collins＆Lanza，2010）。如图1-2所示，两个虚线的分布代表两个潜类别（实线为2个潜类别组成的混合分布）。当潜类别间隔足够大（左图），两个分布的对称轴距离越远（均值差异越大），此时对于来自任一潜类别的个体来说，将其划分到所属类别的精确度越高。如果两个类别间的差异不明显，即潜类别间隔不够大

（右图），两个分布的对称轴距离越近时，将个体精确划分到所属类别就越不容易。这里的潜类别间隔与传统因素分析里的项目区分度或因子负荷的概念类似。当指标只在某个因子上有高负荷，在其他因子上的负荷为0时，将该指标划入某因子是非常确定的。然而如果指标在一个因子上的负荷为0.45，在另外一个因子的负荷为0.48，此时就很难将该指标划入某个单独的因子。显然，潜类别间隔的大小是影响分类精确性的重要因素，也是影响潜类别个数保留的重要变量（Lubke＆Neale，2006）。

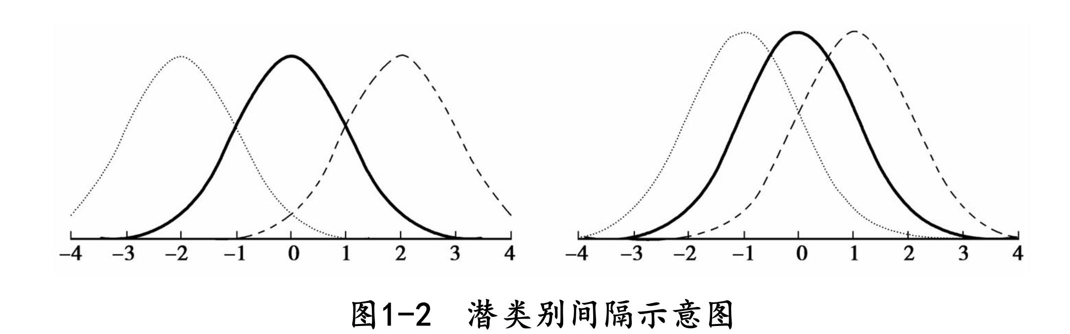
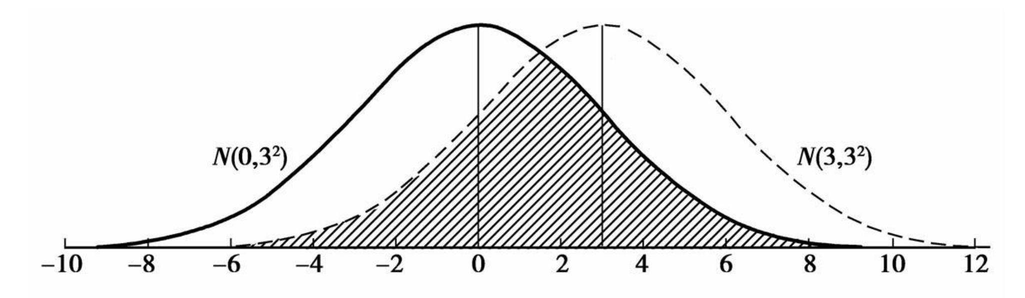
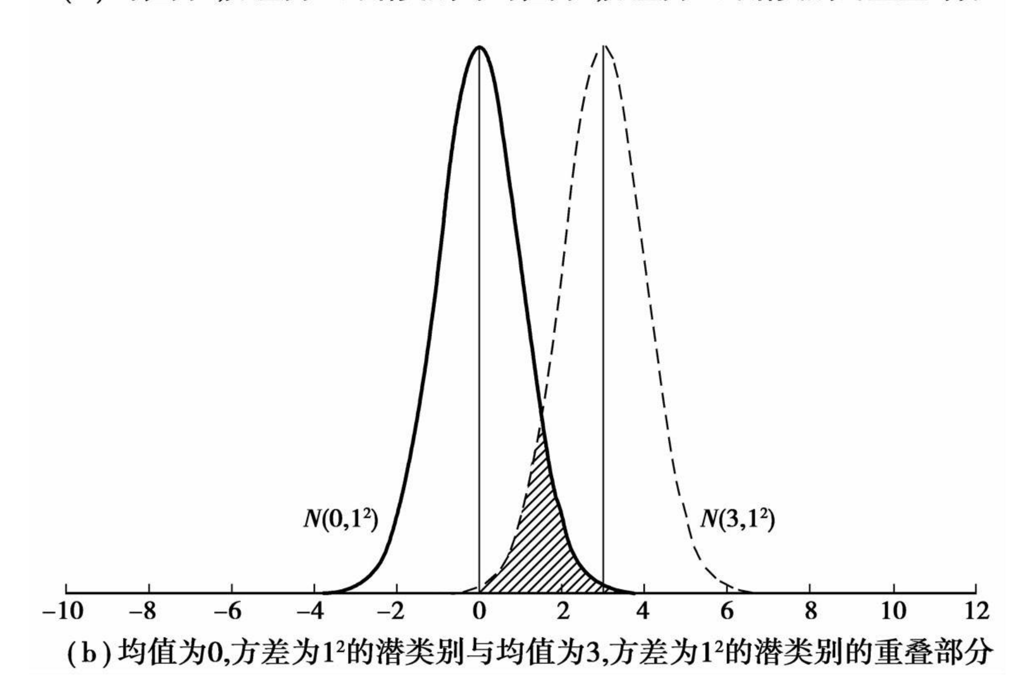


图1-2　潜类别间隔示意图

对分类精确性有影响的另一个因素是类别内方差。在潜类别间距相同的情况下，特定类别分布的方差越大，两个分布之间重叠的部分越大，将个体划分到特定类别组就越困难［图1-3（a）］；类别分布的方差越小，两个分布之间重叠的部分越小，将个体划分到特定类别组就越容易［图1-3（b）］。



(a) 均值为0,方差为3^2的潜类别与均值为3,方差为3^2的潜类别的重叠部分



(b) 均值为0,方差为1^2的潜类别与均值为3,方差为1^2的潜类别的重叠部分

图1-3　类别方差影响分类精确性示意图

1.3.4　参数估计

LCM的参数估计有两种基于极大似然估计的迭代算法：期望最大（Expectation-Maximization，EM）和牛顿－拉夫森算法（Newton-Raphson）。迭代算法的一般过程分成两个阶段：第一阶段，以一个（组）开始值（Starting Values）为起点进行估计以获得最大值；第二阶段，用第一阶段的估计最大值再次进行估计，直到达到设定的聚合标准。

两种方法各有优缺点，但两种算法均容易产生局部最大化解（Local Maxima）而非总体最大化解（Global Maxima）。解决的途径可以是设置不同的起始值估计同一模型，如果结果差异较大说明获得的结果很可能是局部最大化解。在M*plus* 中，程序默认第一阶段从10个随机初始值开始估计；第二阶段，使用第一阶段获得的2个最大值进行估计。当然，研究者可以通过ANALYSIS语句下的STARTS改变这一默认设置，也可以自己设定开始值，具体情况可参见M*plus* 手册例7.4—7.8。

1.4　潜类别分析与其他方法的关系

1.4.1　与因子分析的比较

潜在类别分析和因子分析有很多共同之处，目的都是解释变量间的相关性。在因子模型中，通过抽取公因子使条目之间达成局部独立性；而在LCM中，通过抽取潜类别来实现局部独立性。两者的分析思路是相同的，即通过抽取少数潜变量来解释指标间的关联。然而，两者的区别也是显而易见的，主要体现在以下四方面（Collins＆Lanza，2010）：

第一，潜变量分布形态。因子模型中的潜变量是连续的因子或称维度（Dimension），呈正态分布；而LCM的潜变量是类别的，呈多项式分布（Multinomial Distribution）。

第二，潜变量的性质。维度上的差异存在于量，类别上的差异存在于质（Ruscio＆Ruscio，2008）[1]。

第三，理论取向。LCM假设造成指标关联的原因是不同分布形态的个体混合造成的（如图1-4和图1-5）。图1-4呈现了由两个不同分布形态的正态群体混合后组成一个正态分布的情况。例如，如何解释青少年吸烟和饮酒行为之间的高相关？在FM中，相关是通过一个潜在的“问题行为”（Problem Behavior）因子来解释的（图1-5A）。这个因子是个连续体，存在于所有青少年中，只是程度不同。LCM则认为，两者之间的相关是由于数据存在不同类别群体，一部分（类别一）青少年身上存在高相关，而在另外一部分（类别二）青少年身上相关并不高甚至不存在，如果将这两个群体混合在一起，便观察到吸烟和饮酒间的相关。从此例不难发现，FM的目标是处理变量间的相关性，对象是变量，而LCA的处理对象是个体。

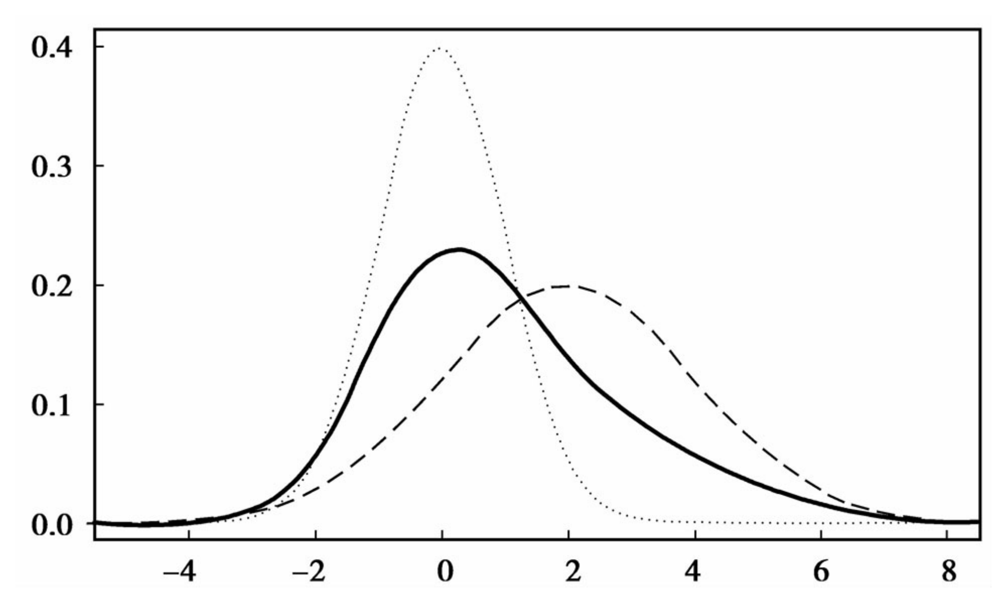


图1-4　两个形态不同的正态样本混合后形成单一的正态分布（实线为联合分布）

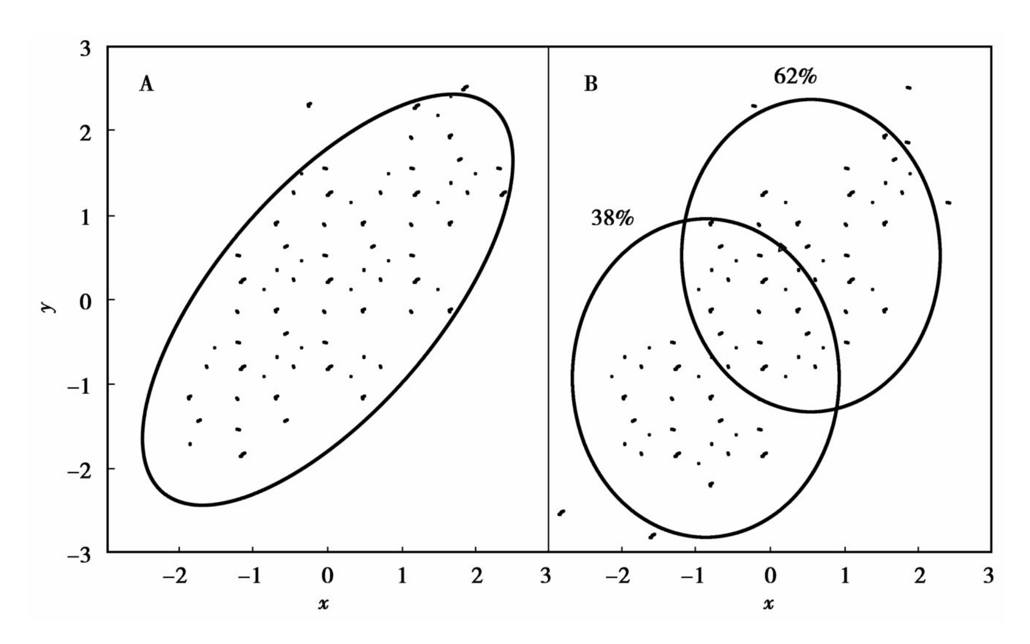


图1-5 FA和LCA用不同的潜变量解释相同的数据（Bauer＆Curran，2004）

第四，分析指标[2]。因子模型中，指标为连续变量，且服从多元正态分布（ML估计下）；而LCM中对变量分布形态的要求比较宽松，可以是任何尺度的变量或不同尺度的组合。分布形态可以是非正态的，但类别内要求正态分布。

更进一步来说，可以根据潜变量和指标的属性将潜类别和因子模型分成如下4类（Bartholomew＆Knott，1999），见表1-4：

表1-4 潜在变量模型分类

潜变量	外显变量	
	类别	连续
类别	潜在类别分析 Latent Class Analysis	潜在剖面分析 Latent Profile Analysis
连续	潜在特质分析/项目反应理论 Latent Trait Analysis/Item Response Theory	因素分析 Factor Analysis

LCM和FM之间的区分更多的是在概念上，统计上两者的界限是非常模糊的。一个抽取*m* 个因子的FA模型可以被*m* ＋1类别的LCM完美地拟合（Bartholomew，1987）。据此，研究者认为两者之间并非相互排斥的关系，而是互为补充（Muthén＆Muthén，2000；Muthén，2003）。在实践中，选择FA还是LCM更多是根据研究的目的，带有很浓的主观色彩。

1.4.2　与聚类分析的异同

LCA作为个体中心（*P* erson-Centered）的方法，与聚类分析在目的上是一致的，即将个体分成不同的群组。由于是基于模型的方法，聚类的标准和结果的检验更合理（Magidson＆Vermunt，2002）。另外，在LCA模型中可以纳入协变量构建回归混合模型，检验更复杂的变量关系（见本书第2章）。

1.4.3　与Taxometric法的异同

Taxometric分析方法最早是由明尼苏达大学Paul Meehl教授及其同事（Meehl＆Yonce，1994；1996）创立和发展起来的。该方法通过探测一组外显变量背后的潜变量是连续的维度还是类别变量来确定群体中是否存在异质群体。按照Taxometric法的术语，一般将异质群体称为类型组（Taxon；与Class类似），剩余的群体称为非类型组或补充组（Complement）。

Taxometric法在功能和目标上与LCA一样，而且同为个体中心的方法，但Taxometric法只将潜在类别变量分成两组（类型组和补足

组），这主要是由于Taxometric法设计之初的目的是用于探测群体中是否存在不同的类别，更确切地说，其关注点是将被试群体区分成两个分群体：类型组（如抑郁症组）和非类型组（健康组）。如果群体中存在两个以上类别的潜在群体，那么Taxometric法将无能为力。McGrath（2008）的分析结果表明，当存在三个类别时，Taxometric法将导致不正确的结论。在类别的本质上，两者也存在差别。Meehl发展Taxometric法的最初目的在于试图证明精神病态和正常人群存在“质”的差别；而LCA将个体分类是基于模型的，并不强调各类别间是否存在本质的差异（Lubke＆Muthén，2005）。Taxometric法不如LCA的另外一点在于不能提供类别概率和单个指标的条件概率。

1.5 潜类别分析的类型

与FM类似，LCM也存在不同的形式。主要有探索性LCM、验证性LCM、多组LCM和多因子LCM。本章只介绍前三种，LCM与其他模型结合的情况将在随后几章单独介绍。

1.5.1 探索性潜类别分析

类似探索性因素分析（EFA）[3]，当事先并不知道潜类别存在几个组别时，需要从实际的数据出发去推测。探索性LCM通常涉及以下四步（邱皓政，2008）：

第一，假定只存在一个类别，即外显变量之间完全独立，也称作零模型或独立模型；

第二，逐步增加潜在类别的个数，计算各模型的参数；

第三，根据拟合评价指标进行模型比较，确定最佳模型；

第四，进行分类和命名。

上述四个步骤为执行探索性LCA的核心步骤，完整的分析还应该包括选择LCA作为分析方法的依据等，结果报告的内容可参见EFA或CFA（王孟成，2014）。

1.5.2　验证性潜类别分析

验证性潜在类别分析（Confirmatory Latent Class Analysis，CLCA）通过增加参数限制来检验特定的假设。在M*plus* 中需要使用参数限定命令。McCutcheon（2002）将参数限制的类型分为如下三种：

（a）等值限定（Equality Restrictions），即检验项目参数组内或跨组相等。多个项目组内等值检验也称平行指标检验（Parallel Indicators），即检验在多个项目上类别内个体具有相同的条件概率或具有相同的误差。跨组等值限定指不同类别在同一项目上具有相同的反应概率。

（b）定值限定（Deterministic Restrictions）即检验类别内或类别间在某些项目上的条件概率等于特定的值，如假定类别一在第2个项目上的反应概率为0或1。

（c）不等值限定（Inequality Restrictions）。即检验不同类别在项目上反应概率不等值的假设是否成立。一般包括非线性不等值

和顺序关系两种（Finch＆Bronk，2011）。

1.5.3　多组潜类别分析

多组LCA模型是LCA模型向多组的扩展。同多组CFA一样，有时根据研究需要对多个群体或时间点的潜类别结构进行检验，因此也称同时潜类别分析（Simultaneous LCA）。完全不限定的多组LCA被称作完全异质模型（Complete Heterogeneity），即用LCA模型分别估计各组数据，类似MCFA时的未设限的形态等值模型。

多组LCA的参数限制涉及两部分：条件概率等值和潜类别概率等值，也称作测量等值（Measurement Equivalence）和类别大小等值（Class Sizes Equivalence；Eid，Langeheine＆Diener，2003）。

1.5.4　多个潜变量的潜类别模型（Latent Class Factor Model，LCFM）[(4)]

上述三种LCA形式均假设指标只被单一的潜类别变量解释，类似单维性（Unidimensionality）假设。而LCFM假设存在两个以上的潜类别变量共同影响指标的关系，如图1-6所示，C_a 和C_b 为两个潜类别变量（可以有不同的水平）共同影响着5个指标。

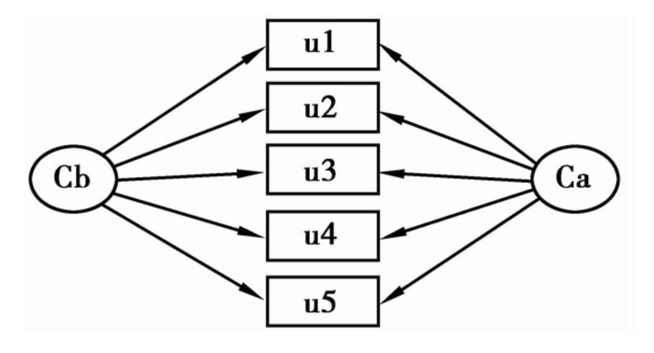


图1-6　多潜变量LCFM示意图

LCFM在实践中很少使用，因此本章不做探讨。关于LCFM更多的信息请参见（Magidson＆Vermunt，2001；邱皓政，2008）。

1.5.5　包含协变量的潜类别分析

不包含协变量的LCA也称作无条件LCA，包含协变量的LCA也称作回归混合模型（Regression Mixture Modeling，RMM；e. g. ，Clark＆Muthén，2009），将在第2章详细介绍。

1.5.6　其他形式

除了上述几种常见的LCA形式，LCM还可以拓展到分析具有嵌套数据结构的多水平LCA（Multilevel Latent Class Analysis，MLCA；Vermunt，2003）。具体的实例如Henry和Muthén（2010）采用MLCM分析美国妇女吸烟状况的研究。

1.6　模型拟合评价 [(5)]

确定潜类别数目是混合模型（Mixture Model）需要考虑的重要问题，低估和高估类别数据均不理想。目前，用于确定类别数目的指标可以分成两类：信息指数和基于似然比的检验统计量。

1.6.1 模型评价

模型的适配检验方法主要有Peason卡方检验和似然比卡方G^2（LL）检验，以及信息评价指标AIC、BIC和样本校正的BIC（sample size-adjusted BIC，aBIC）。这几种统计量都是通过比较期望值与实际值差异来判断模型拟合的优劣，统计值越小表示拟合得越好。然而，当样本量很大时，卡方统计检验变得十分保守，即使期望值与观察值相差不大，也很容易判断为差异显著。所以大部分的实证研究都使用BIC指标作为模型适配度比较的指标，一般是选择BIC最小的模型作为最佳模型（Xian et al，2005）。

$$\mathrm{aBIC}_1 = -2\log L + t\log((N+2)/24) \tag{1.10}$$

在评价LCA时，还常常使用Entropy去评价分类精确性，公式如下：

$$E_k = 1 - \frac{\sum_i \sum_k (-p_{ik}\ln p_{ik})}{n\ln K} \tag{1.11}$$

P_{ik} 为个体i 属于类别k 的后验概率。其取值范围为0～1，越接近1表明分类越精确。Lubke和Muthén（2007）指出，Entropy＜0.60相当于超过20％的个体存在分类错误；Entropy＝0.80表明分类准确率超过90％。然而最近的模拟研究发现（王孟成等，2017）：①尽管

Entropy值与分类精确性高相关，但其值随类别数、样本量和指标数的变化而变化，很难确定唯一的临界值；②其他条件不变的情况下，样本量越大，Entropy的值越小，分类精确性越差；③类别距离对分类精确性的影响具有跨样本量和跨类别数的一致性；④小样本（N＝50～100）的情况下，指标数越多，Entropy的结果越好；⑤在各种条件下Entropy对分类错误率比其他变式更灵敏。

1.6.2　模型比较

似然比检验（LRT）并不适合嵌套LCA模型间的差异比较，主要由于k －1个类别是k 个类别模型的特例，在k 个类别模型中，第一个类别的概率被设定为0，导致差异间不再是卡方分布。所以替代的检验统计量被提出，其中基于Bootstrap的似然比检验（BLRT；McLachlan＆Peel，2000）和LMR（Lo-Mendell-Rubin LMR；Lo，Mendell＆Rubin，2001）亦称作aLRT（Adjusted LRT），较为流行。

BLRT使用Bootstrap抽样估计两个嵌套模型间的对数似然比差异分布。BLRT主要比较k －1个和k 个类别模型间的拟合差异。例如，对于一个有4个类别的LCA模型，BLRT的p 值比较3个类别和4个类别模型间拟合的差异。显著的BLRT p 值表示4个类别的模型比3个类别的模型拟合显著改善。不显著的BLRT p 值则表明4个类别的模型并未比三个类别模型显著改善拟合。BLRT值在M*plus* 中通过TECH14获得，也可以通过在ANALYSIS命令下调用LRTBOOTSTRAP获得更精确的估计（Muthén＆Muthén，2010），通常推荐的BOOTSTRAP值为100（McLachlan＆Peel，2000），增加BOOTSTRAP值会延长计算时间。

LMR与BLRT类似，用于比较k －1个和k 个类别模型间的拟合差异。显著的LMR p 值表明k 个类别模型优于k －1个类别模型。M*plus*除了提供LMR还提供LMR的校正值，只是比较所使用的“k －1模型”为删除第一个类别的模型（Muthén＆Muthén，2010）。

1.6.3　指数评价——保留正确的类别个数

保留正确的潜类别个数是目前LCA乃至混合模型领域的热点和难点问题。高估和低估潜类别个数都将影响结果推论的准确性，不少模拟研究对这一问题进行了探索（e．g．，Dias，2006；Nylund et al，2007；Yang，2006）。

模拟研究发现（Yang，2006），aBIC是分类准确度最高的信息指数，其前提是每个类别至少要有50个被试（Yang，2006）。例如，潜类别数有5个，样本量至少要有250个才能保证准确性。Yang的研究还发现，AIC在确定类别个数时表现欠佳（也见，Nylund et al，2007），但也有模拟研究发现AIC3表现最佳（Dias，2007）。Nylund等（2007）的研究在多种混合模型（LCA，因子混合模型和增长混合模型）中比较了BLRT，LMR和信息指数，结果发现BLRT和BIC分别为基于似然比和信息指数中表现最好的指标，而BLRT又优于BIC，且在多种情况下（样本量、模型类型、指标个数及潜类别个数）具有一致的表现。

在实际应用中，各评价指标之间并不一致。例如，BLRT的p 值显著，而LMR的p 值则远大于0。如果遇到此种情况，应结合分类的实际

意义和类别所包含的样本数来确定最终的类别数目。具体来说，即使各项指标提示保留m个类别，而其中的一个类别个体数目有限或者不易解释时，应该考虑$m-1$个类别的模型。

通常似然值或信息指数会随着类别数目的增加而减少，但有时类别数目增加了很多也未必能获得最佳的拟合模型。例如，BIC值随类别数单调递减，始终未见最低值。遇到这种情况时，可以采用类似EFA中确定因子个数的陡坡图检验（Petras＆Masyn，2010）。例如，图1-7就是根据表1-5（本书23页）中的aBIC值从高到低依次排列的陡坡图。仅从图上来看，在2处存在明显的拐点，因此选取2个类别是合适的。

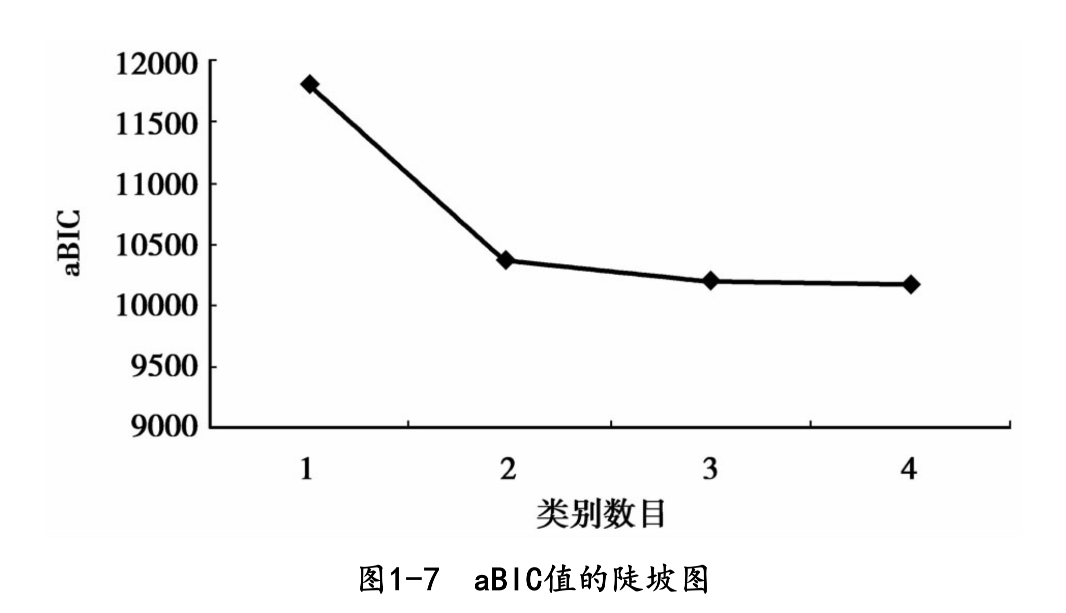


图1-7　aBIC值的陡坡图

1.6.4　潜类别分析的其他问题

LCA作为个体中心化的方法，通过对个体在项目反应上的列联表分析将群体分成若干潜在组别。模型检验时的皮尔逊卡方（Pearson Chi-Squared）和似然比（Likelihood-Ratio）检验对列联表单元格的期望次数有一定的要求（例如，在5个以上），否则影响检验精确性。

在样本量一定的前提下，随着条目数量的增加，单元格数目随之增加，每个单元格的期望次数随之减少。所以做LCA时对样本量有一定要求，Yang（2006）从模型分类准确性的角度出发，发现每个类别至少要有50个被试才能保证根据aBIC正确选择模型（Yang，2006）。目前研究对LCA样本量影响模型选择和参数估计的研究并不多，将来的研究应会对此问题做进一步探索。

1.7 探索性潜类别分析示例

1.7.1 探索性LCA的M*plus* 分析过程

为了演示标准LCA的分析过程，我们将PTSD的数据转化成0，1计分。将原始计分小于3分的记0分，大于或等于3分的记1分。这样做的依据是根据PCL-C的计分特点，得分大于3作为阳性症状。框1-1中呈现了抽取三个潜类别的M*plus* 程序及注释，框1-2是对应的输出结果。

框1-1 探索性LCA的M*plus* 语句

```
TITLE: This is an example of a classic LCA
DATA: FILE IS PTSD__2. dat;
VARIABLE: NAMES ARE x y0 y1-y17;
   USEVARIABLES ARE y1-y17;
   CATEGORICAL＝y1-y17;
   CLASSES＝c（3）；！设定潜类别个数，从1个类别开始，依次增加;
   Auxiliary＝x y0；！设置辅助变量;
```

```
ANALYSIS：TYPE＝MIXTURE；
    STARTS＝200 50；！避免局部最大化解，增加随机起始值数；
OUTPUT：TECH11 TECH14；
SAVEDATA：FILE＝ptsdlca-1. txt；
    Save＝cprob；！保存后验分类概率；
    Plot：！通过绘图命令，可以获得描述性统计图和条件概率示意图。
    type is plot3；
    series＝y1-y17（*

）；
```

框1-2　三类别探索性LCA分析M*plus* 的结果（有删减）

```
! 下面是被试在 17 个条目的 2 个选项上的回答比例和计数。
UNIVARIATE PROPORTIONS AND COUNTS FOR CATEGORICAL VARIABLES
    Y1
      Category 1      0.629        352.000
      Category 2      0.371        208.000
    .
.
    Y17
      Category 1      0.486        272.000
      Category 2      0.514        288.000
RANDOM STARTS RESULTS RANKED FROM THE BEST TO THE WORST
LOGLIKELIHOOD VALUES
```

```
! 从 200 个初始值中确定的 10 个最优值使得似然值最大。
Final stage loglikelihood values at local maxima, seeds, and initial stage start numbers:
          -5014.461   695155          150
          -5014.461   609185          181
          -5014.462   804561          59
          -5014.465   626891          32
          -5014.465   715561          125
          -5014.466   347515          24
          -5014.466   915107          54
          -5014.466   533738          11
          -5014.466   unperturbed     0
          -5014.466   354559          73
WARNING: WHEN ESTIMATING A MODEL WITH MORE THAN TWO CLASSES, IT MAY BE
NECESSARY TO INCREASE THE NUMBER OF RANDOM STARTS USING THE STARTS OPTION TO
AVOID LOCAL MAXIMA. ! 提示增加起始值数的警告信息
THE MODEL ESTIMATION TERMINATED NORMALLY
MODEL FIT INFORMATION
Number of Free Parameters                             53
Loglikelihood
          H0 Value                               -5014.461
          H0 Scaling Correction Factor               1.091
            for MLR
Information Criteria
          Akaike (AIC)                           10134.922
          Bayesian (BIC)                         10364.303
          Sample-Size Adjusted BIC               10196.055
          (n* = (n + 2) / 24)
Chi-Square Test of Model Fit for the Binary and Ordered Categorical
(Ordinal) Outcomes**
          Pearson Chi-Square
          Value                                   6898.152
          Degrees of Freedom                        130794
          P-Value                                   1.0000
              Likelihood Ratio Chi-Square
          Value                                   1205.988
          Degrees of Freedom                        130794
          P-Value                                   1.0000
** Of the 131072 cells in the latent class indicator table, 224 were deleted in the calculation of chi-
```

square due to extreme values.

FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES

BASED ON THE ESTIMATED MODEL(模型估计的潜类别概率)

Latent Classes		
1	134.54385	0.24026
2	240.78221	0.42997
3	184.67394	0.32977

FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASS PATTERNS

BASED ON ESTIMATED POSTERIOR PROBABILITIES

Latent Classes		
1	134.48299	0.24015
2	240.91948	0.43021
3	184.59753	0.32964

CLASSIFICATION QUALITY(分类精确性指标 ***Entropy***)

Entropy 0.801

CLASSIFICATION OF INDIVIDUALS BASED ON THEIR MOST LIKELY LATENT CLASS MEMBERSHIP(分类结果,每个类别的个数和比例)

Class Counts and Proportions

Latent Classes		
1	128	0.22857
2	247	0.44107
3	185	0.33036

Average Latent Class Probabilities for Most Likely Latent Class Membership (Row) by Latent Class (Column)后验概率的平均值

	1	2	3
1	0.918	0.082	0.000
2	0.069	0.884	0.047
3	0.000	0.065	0.935

MODEL RESULTS(每个潜类别下的指标阈限值,空间限制只显示前三个指标)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Latent Class 1				
Thresholds				
Y1 $ 1	-0.466	0.211	-2.211	0.027
Y2 $ 1	-0.109	0.287	-0.381	0.703
Y3 $ 1	-1.010	0.235	-4.306	0.000
Latent Class 2				
Thresholds				
Y1 $ 1	0.247	0.183	1.351	0.177
Y2 $ 1	0.950	0.169	5.608	0.000
Y3 $ 1	0.380	0.257	1.477	0.140
Latent Class 3				
Thresholds				
Y1 $ 1	2.123	0.346	6.129	0.000
Y2 $ 1	2.980	0.481	6.191	0.000
Y3 $ 1	3.024	0.549	5.505	0.000
Categorical Latent Variables				
Means				
C#1	-0.317	0.309	-1.024	0.306
C#2	0.265	0.150	1.772	0.076

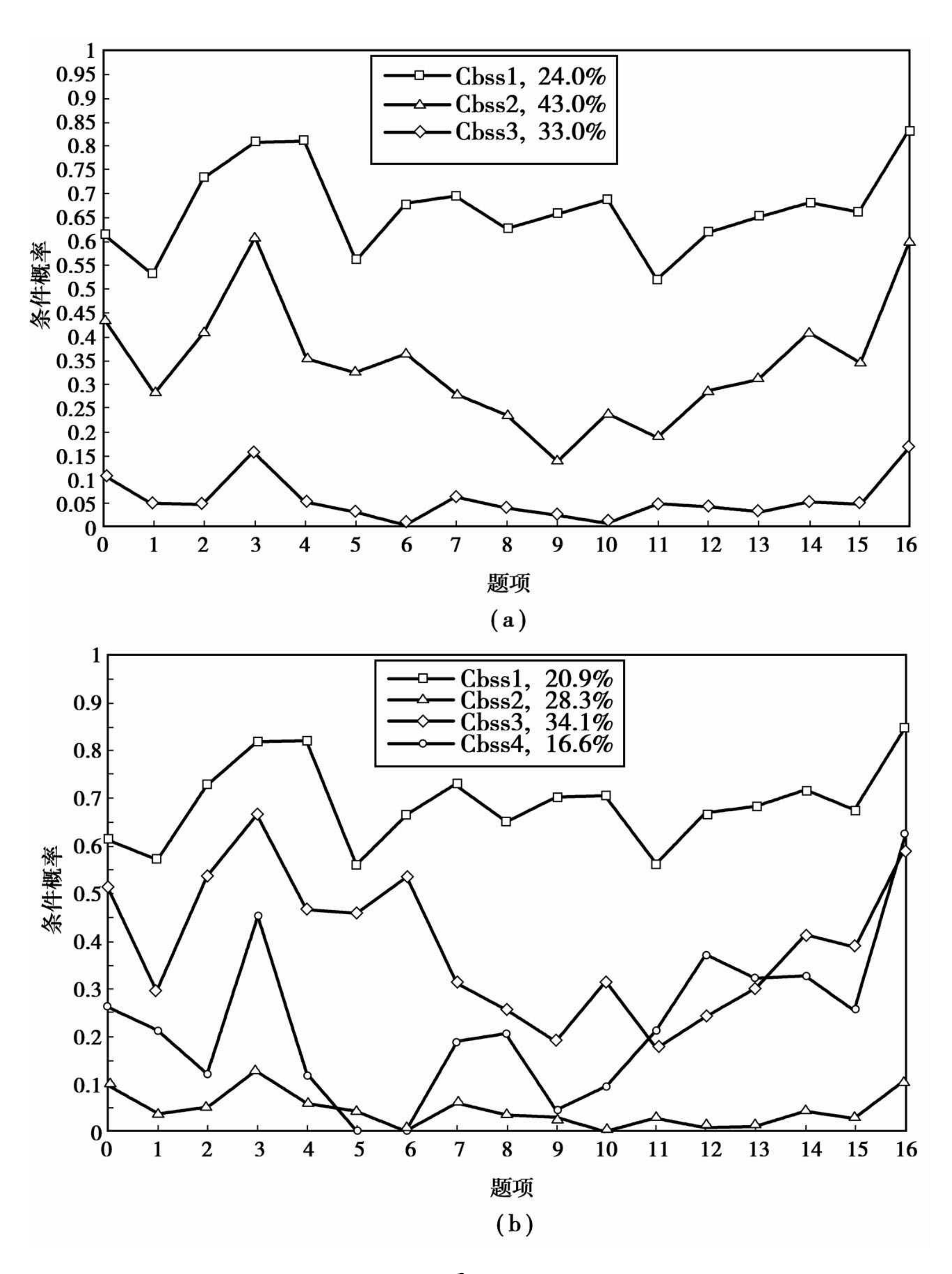
RESULTS IN PROBABILITY SCALE(条件概率,与图 1-8 结果一致)

```
Latent Class 1
Y1
      Category 1        0.386      0.050      7.729      0.000
      Category 2        0.614      0.050     12.311      0.000
.
.
.
Y17
      Category 1        0.167      0.047      3.583      0.000
      Category 2        0.833      0.047     17.916      0.000
Latent Class 2
Y1
      Category 1        0.561      0.045     12.472      0.000
      Category 2        0.439      0.045      9.743      0.000
.
.
Y17
      Category 1        0.399      0.041      9.699      0.000
      Category 2        0.601      0.041     14.610      0.000
Latent Class 3
Y1
      Category 1        0.893      0.033     27.006      0.000
      Category 2        0.107      0.033      3.233      0.001
.
.
Y17
      Category 1        0.831      0.049     17.070      0.000
      Category 2        0.169      0.049      3.465      0.001
LATENT CLASS ODDS RATIO RESULTS(为节省版面只显示前三个变量的结果)
Latent Class 1 Compared to Latent Class 2
Y1
      Category > 1      2.039      0.567      3.595      0.000
Y2
      Category > 1      2.885      0.941      3.066      0.002
Y3
      Category > 1      4.013      1.334      3.009      0.003
Latent Class 1 Compared to Latent Class 3
Y1
      Category > 1     13.305      4.998      2.662      0.008
Y2
      Category > 1     21.959     10.445      2.102      0.036
Y3
      Category > 1     56.463     32.052      1.762      0.078
```

```
Latent Class 2 Compared to Latent Class 3
Y1
    Category > 1        6.524       2.344       2.783       0.005
Y2
    Category > 1        7.612       3.853       1.975       0.048
Y3
    Category > 1       14.069       7.354       1.913       0.056
QUALITY OF NUMERICAL RESULTS
     Condition Number for the Information Matrix              0.475E-02
        (ratio of smallest to largest eigenvalue)
TECHNICAL 11 OUTPUT (LMR 检验的信息)
     Random Starts Specifications for the k-1 Class Analysis Model
          Number of initial stage random starts                  200
          Number of final stage optimizations                     10
VUONG-LO-MENDELL-RUBIN LIKELIHOOD RATIO TEST FOR 2 (H0) VERSUS 3 CLASSES
             H0 Loglikelihood Value                        -5130.699
             2 Times the Loglikelihood Difference           232.475
             Difference in the Number of Parameters              18
             Mean                                            29.407
             Standard Deviation                              41.112
             P-Value                                         0.0018
     LO-MENDELL-RUBIN ADJUSTED LRT TEST
             Value                                          230.452
          P-Value                                            0.0019
TECHNICAL 14 OUTPUT(BLRT 检验信息)
     Random Starts Specifications for the k-1 Class Analysis Model
          Number of initial stage random starts               200
          Number of final stage optimizations                  10
           Random Starts Specification for the k-1 Class Model for Generated Data
          Number of initial stage random starts                   0
          Number of final stage optimizations for the
             initial stage random starts                          0
     Random Starts Specification for the k Class Model for Generated Data
          Number of initial stage random starts                  20
          Number of final stage optimizations                     5
     Number of bootstrap draws requested                     Varies
PARAMETRIC BOOTSTRAPPED LIKELIHOOD RATIO TEST FOR 2 (H0) VERSUS 3 CLASSES
(BLRT 检验结果)
             H0 Loglikelihood Value                        -5130.699
             2 Times the Loglikelihood Difference           232.475
             Difference in the Number of Parameters              18
             Approximate P-Value                             0.0000
             Successful Bootstrap Draws                           5
```

```
PLOT INFORMATION(图形信息)
The following plots are available:
  Histograms (sample values)
  Scatterplots (sample values)
  Sample proportions
  Estimated probabilities
SAVEDATA INFORMATION(数据保存的信息)
  Order and format of variables
    Y1              F10.3
   .
   .
   .
    Y17             F10.3
    X               F10.3
    Y0              F10.3
    CPROB1          F10.3 ! 后验分类概率
CPROB2              F10.3
    CPROB3          F10.3
    C               F10.3 ! 每个样本所属类别的标签。
(保存的文件名,格式和长度)
Save file ptsd-lca-3.txt
Save file format   23F10.3
Save file record length    5000
```

通过PLOT命令可以获得条件概率图，在运行结束后可通过GRAPH的下拉菜单查看。*Mplus* 会提供在每个类别上的条件概率折线图，如图1-8和图1-9所示。



Cbss1, 24.0%
Cbss2, 43.0%
Cbss3, 33.0%
条件概率
题项
(a)
Cbss1, 20.9%
Cbss2, 28.3%
Cbss3, 34.1%
Cbss4, 16.6%
条件概率
题项
(b)

图1-8

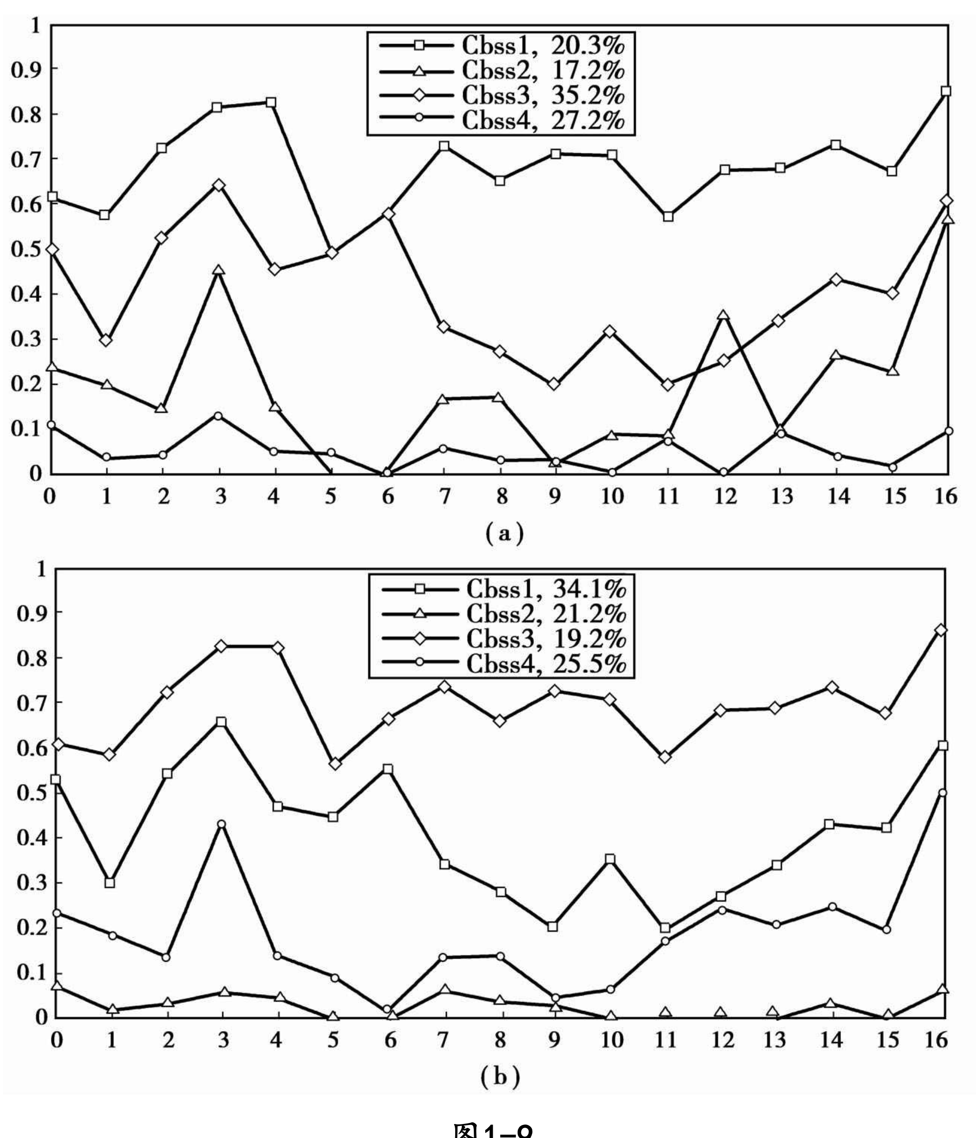


图1-9

1.7.2 结果解释

抽取1～5个类别的拟合指数汇总在表1-5中。具体来说，信息指数提示的结论并不一致，保留3个类别时的BIC值最小，保留4个类别时的aBIC值最小，保留5个类别时的AIC值最小。而保留2个类别时的G^2值和卡方值最小。LMR和BLRT的结果也不一致，保留4个类别时LMR不显著，支持保留3个类别，而BLRT则支持保留4个类别。任何时候，统计

指标只是提供决策参考，在确定最佳模型时还应该考虑各类别的可解释性。从图1-8（a）和（b）可见，保留3个和4个类别时，第一个和最后一个类别的条件概率非常的一致（两个类别非常稳定）。在4类别模型中的第3和第4类是由3类别模型中的第2个类别分离出来的。具体来看，4类别模型中的第3和第4类在条目3，6，7，10和11上的条件概率差异明显，所以综合考虑后选择4类别模型为最佳模型。需要注意的是，4类别模型的Entropy最小，提示可能存在分类不精确的问题。

表1-5　潜类别分析拟合信息汇总表

Model	k	G^2/LL	*df*	χ^2	AIC	BIC	aBIC	Entropy	LMR	BLRT	类别概率
1C	17	2234.99	130832	17316.34	11780.14	11853.72	11799.75	—	—	—	1
2C	35	**789.26**	130753	**4478.20**	10331.40	10482.88	10371.77	**.865**	<.0001	<.0001	.512/.488
3C	53	1205.99	130794	6898.15	10134.92	**10364.30**	10196.06	.801	.0019	<.0001	.229/.441/.330
4C	71	1161.34	130778	6392.11	10101.33	10408.62	**10183.23**	.769	.6314	<.0001	.205/.295/.336/.164
5C	89	1121.51	130761	6447.53	**10089.17**	10474.36	10191.83	.775	.2621	.066	.191/.336/.109/.120/.245
CLCA											
等值	67	1088.96	130771	5817.05	10136.45	10426.42	10213.73	.758	.579	<.0001	.198/.164/.350/.288
定值	65	1178.60	130785	6772.32	10099.80	10381.11	10174.77	.755	.2173	<.0001	.341/.234/.189/.236
不等值	70	1161.48	130779	6392.658	10099.41	10402.36	10180.15	.769	_a	<.0001	.295/.166/.205/.334

注：最佳值被加粗；a＝非线性限制时不可得。

根据4个类别在17个条目上的应答特点，可将其命名为高症状组（115人），低症状组（165人），情感麻木组（188人，项目8-12主要反映麻木症状）和低回避组（92人，项目6和7反映的是回避症状）。

1.8　潜在剖面分析示例

在实践中，分析的变量通常是连续型的，所以潜在剖面分析更为常见。下面我们将PTSD的5级计分当作连续型变量来演示LPA的分析过程。

LPA的分析语句和LCA基本一致，唯一的区别在于将分析指标当作连续变量，体现在语句上即将原LCA指定变量类型的“CATEGORICAL”去掉。框1-3呈现了LPA的M*plus* 语句。

框1-3　LPA示例语句

```
TITLE: This is an example of a classic LCA
DATA: FILE IS PTSD_2.dat;
VARIABLE: NAMES ARE x y0 y1-y17;
          USEVARIABLES ARE y1-y17;
          CLASSES = c (3);
ANALYSIS: TYPE = MIXTURE;
          STARTS = 20050;
OUTPUT: TECH11 TECH14;
```

```
SAVEDATA: FILE = ptsdlca-1.txt;
Save = cprob;
Plot: type is plot3;
series = y1-y17(*);
```

```
MODEL RESULTS
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Latent Class 1
Means
   Y1                  1.823      0.060     30.381      0.000
   Y2                  1.473      0.053     27.823      0.000
   Y3                  1.529      0.058     26.425      0.000
   Y4                  2.075      0.068     30.318      0.000
   Y5                  1.526      0.058     26.527      0.000
Variances
   Y1                  1.048      0.066     15.973      0.000
   Y2                  1.114      0.074     15.080      0.000
   Y3                  0.956      0.064     14.853      0.000
   Y4                  1.194      0.073     16.440      0.000
   Y5                  1.110      0.080     13.923      0.000
Latent Class 2
Means
   Y1                  2.747      0.090     30.390      0.000
   Y2                  2.293      0.102     22.545      0.000
   Y3                  2.787      0.095     29.260      0.000
   Y4                  3.098      0.101     30.545      0.000
   Y5                  2.722      0.114     23.779      0.000
Variances
   Y1                  1.048      0.066     15.973      0.000
   Y2                  1.114      0.074     15.080      0.000
   Y3                  0.956      0.064     14.853      0.000
   Y4                  1.194      0.073     16.440      0.000
```

Y5	1.110	0.080	13.923	0.000

在结果输出上也存在一些差别。图1-10的上图是LCA的结果图示选项，下图是LPA的结果选项。相比较而言，LPA输出的图示结果较丰富些。如图1-10所示，LCA输出的是条件概率图，而LPA输出的则是条件均值图，图示结果见图1-11。另外，LPA类别数目的保留同LCA。

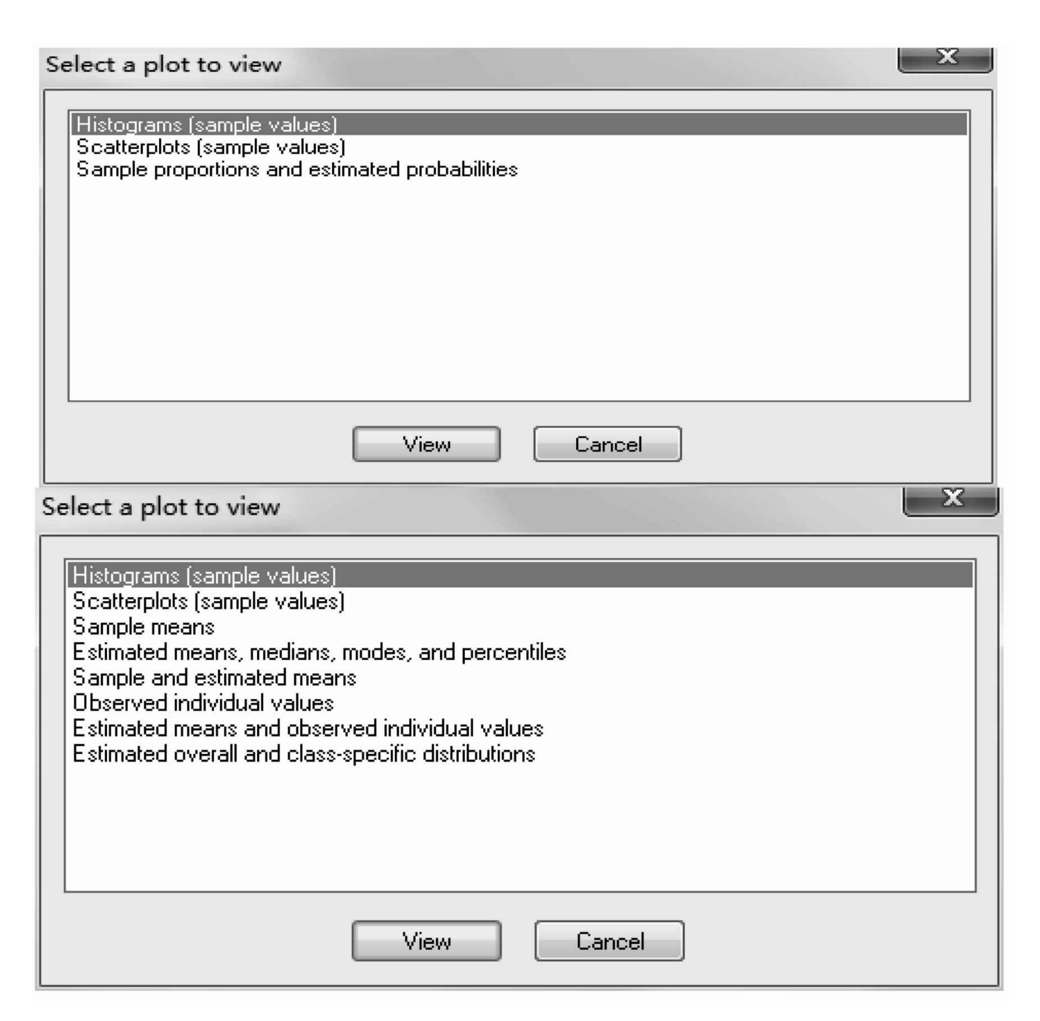


图1-10　view plot选项对话框（上图LCA下图LPA）

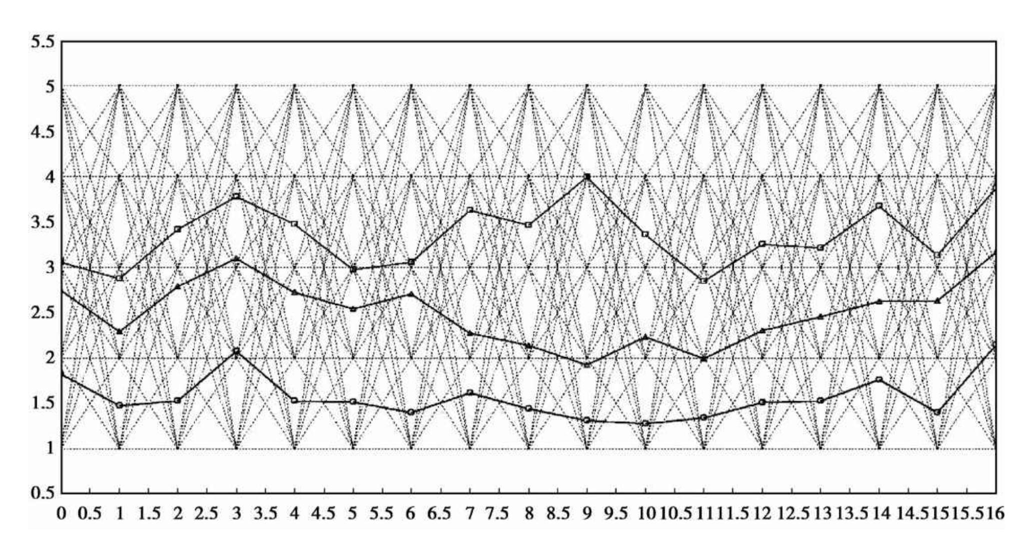


图1-11　LPA三个类别估计条件均值和个体值

1.9　验证性潜类别分析示例

在M*plus* 中，限制参数是通过限制二分变量阈限值来实现的。反应概率和指标阈限值之间存在如下关系（Muthén，2001）：

$$p = \frac{1}{1 + e^{-\tau}} \tag{1.12}$$

正的阈限值越大表明特定类别反应概率越小，而大的负值表明反应概率越大。例如，＋3表示非常低的特定反应概率，对应概率值为0.047，－3代表一个较高的反应概率，对应概率值为0.953。通常，阈限值＋1和－1分别代表低概率和高概率。

下面以PTSD数据为例分别假设三种限制条件：

（a）等值限定。根据探索性分析的结果或理论等假设信息，假定C1和C3在条目6和7上有相等的条件概率即等值；C2和C4在条目12和14

上的条件概率等值。M*plus* 语句和部分结果见框1-4。从框中结果部分可见，对应项目的阈限值是相等的。

框1-4　等值限定的M*plus* 语句和部分结果

```
TITLE: CLCA for 4 classes of PTSD
DATA: FILE IS PTSD_2.dat;
VARIABLE: NAMES ARE x y0 y1-y17;
          USEVARIABLES ARE y1-y17;
          CATEGORICAL =y1-y17;
CLASSES = c (4);
ANALYSIS: TYPE = MIXTURE;
STARTS=20010;
LRTBOOTSTRAP=100;
MODEL: %overall%
%c#1%
[y6$1*3] (1);
[y7$1*3] (2);
%c#2%
[y12$1*-2] (3);
[y14$1*-2] (4);
%c#3%
[y6$1*3] (1);
[y7$1*3] (2);
%c#4%
[y12$1*-2] (3);
[y14$1*-2] (4);
OUTPUT: TECH11 TECH14;
```

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Latent Class 1				
Thresholds				
Y1$1	-0.451	0.233	-1.935	0.053
Y6$1	0.038	0.153	0.248	0.804
Y7$1	-0.321	0.167	-1.917	0.055
Y12$1	-0.269	0.260	-1.034	0.301
Y13$1	-0.732	0.345	-2.121	0.034
Y14$1	-0.749	0.264	-2.840	0.005
Latent Class 2				
Thresholds				
Y1$1	2.120	0.846	2.505	0.012
Y6$1	3.077	0.502	6.131	0.000
Y7$1	15.000	0.000	999.000	999.000
Y12$1	[illegible]	[illegible]	[illegible]	[illegible]

Y12 $ 1	2.416	0.281	8.609	0.000
Y13 $ 1	8.982	115.173	0.078	0.938
Y14 $ 1	2.270	0.446	5.094	0.000
Latent Class 3				
Thresholds				
Y1 $ 1	−0.004	0.226	−0.016	0.987
Y6 $ 1	0.038	0.153	0.248	0.804
Y7 $ 1	−0.321	0.167	−1.917	0.055
Y12 $ 1	1.425	0.263	5.409	0.000
Y13 $ 1	1.098	0.254	4.328	0.000
Y14 $ 1	0.651	0.335	1.944	0.052
Latent Class 4				
Thresholds				
Y1 $ 1	1.148	0.928	1.237	0.216
Y6 $ 1	15.000	0.000	999.000	999.000
Y7 $ 1	15.000	0.000	999.000	999.000
Y12 $ 1	2.416	0.281	8.609	0.000
Y13 $ 1	0.642	0.972	0.660	0.509
Y14 $ 1	2.270	0.446	5.094	0.000

注：限定相等的参数用相同符号标记。

模型拟合参数结果呈现在表1-7。由于限制等值的模型嵌套于不受限模型，所以通过嵌套模型似然比检验来比较增加等值限定是否恶化了模型拟合。限定等值的模型自由度比未限定模型自由度减少了7个（130778－130771＝7），相应的G^2值不但没有增加反而下降了72.38（$p<0.01$），仅从拟合的角度考虑，等值限制是可取的。

（b）定值限定即检验项目条件概率等于特定的值，如0或1。假定C2组在条目6，7和1－14上的条件概率[6]为0，即设置阈限值为15（根据阈限值与概率间关系的公式，$\tau=15$时，p非常接近于0），其他组自由估计。

从框1-5的结果可见，C2的6个项目的阈限值固定为15，对应的概率为1，其他类别自由估计。定值的模型自由度比未限定模型自由度增加了7个（130785－130778＝7），相应的G^2值增加了17.3（$p < 0.01$），仅从拟合的角度考虑，定值限制恶化了模型拟合，即定值限制不合理。

框1-5　定值限定的M*plus* 语句和部分结果

```
    MODEL: %overall%
    %c#2%
    [y6$1@15];! 通过@符号设置固定值;
    [y7$1@15];
    [y11$1@15];
    [y12$1@15];
    [y13$1@15];
    [y14$1@15];
Latent Class 1
  Thresholds
    Y6$1            -0.255      0.252     -1.012      0.312
    Y7$1            -0.685      0.270     -2.543      0.011
    Y8$1            -1.042      0.305     -3.413      0.001
    Y9$1            -0.659      0.285     -2.317      0.021
    Y10$1           -0.974      0.447     -2.179      0.029
    Y11$1           -0.897      0.281     -3.194      0.001
    Y12$1           -0.304      0.288     -1.055      0.291
    Y13$1           -0.771      0.388     -1.987      0.047
    Y14$1           -0.794      0.283     -2.804      0.005
  Latent Class 2
  Thresholds
    Y6$1            15.000      0.000    999.000    999.000
    Y7$1            15.000      0.000    999.000    999.000
    Y8$1             2.729      0.555      4.914      0.000
    Y9$1             3.286      0.602      5.460      0.000
    Y10$1            3.498      0.665      5.258      0.000
    Y11$1           15.000      0.000    999.000    999.000
    Y12$1           15.000      0.000    999.000    999.000
    Y13$1           15.000      0.000    999.000    999.000
    Y14$1           15.000      0.000    999.000    999.000

RESULTS IN PROBABILITY SCALE
  Latent Class 1
  Y6
    Category 1       0.553      0.050     10.980      0.000
    Category 2       0.447      0.050      8.886      0.000
  Y7
    Category 1       0.442      0.064      6.909      0.000
    Category 2       0.558      0.064      8.736      0.000
```

```
  Y11
    Category 1       0.651      0.073      8.955      0.000
    Category 2       0.349      0.073      4.801      0.000
  Y12
    Category 1       0.809      0.056     14.452      0.000
    Category 2       0.191      0.056      3.419      0.001
```

Y13				
Category 1	0.729	0.057	12.761	0.000
Category 2	0.271	0.057	4.741	0.000
Y14				
Category 1	0.663	0.079	8.421	0.000
Category 2	0.337	0.079	4.281	0.000
Latent Class 2				
Y6				
Category 1	1.000	0.000	0.000	1.000
Category 2	**0.000**	0.000	0.000	1.000
Y7				
Category 1	1.000	0.000	0.000	1.000
Category 2	**0.000**	0.000	0.000	1.000
Y11				
Category 1	1.000	0.000	0.000	1.000
Category 2	**0.000**	0.000	0.000	1.000
Y12				
Category 1	1.000	0.000	0.000	1.000
Category 2	**0.000**	0.000	0.000	1.000
Y13				
Category 1	1.000	0.000	0.000	1.000
Category 2	**0.000**	0.000	0.000	1.000
Y14				
Category 1	1.000	0.000	0.000	1.000
Category 2	**0.000**	0.000	0.000	1.000

（c）不等值限定。即通过不等值限制检验不同类别在项目上反应概率是否存在差异。有时根据理论或先前的研究结果假定不同类别间在某些条目上的条件概率存在大小顺序，或其他不等值关系时，可以使用不等价限定来实现。假定C1在项目7上的条件概率是C2的2倍；C3在项目7上概率比C4概率大。从框1-6的结果可见，在项目7上，C1的阈限值为36.977，C2的阈限值为18.488。

框1-6　不等值限定M*plus* 语句和部分结果

```
MODEL: %overall%
%c#1%
[y7$1] (p1);
  %c#2%
  [y7$1] (p2);
  %c#3%
  [y7$1*3] (p3);
  %c#4%
  [y7$1*3](p4) ;
  MODEL CONSTRAINT:
    P1=2*p2;
    P4>p3;
OUTPUT: TECH11 TECH14;
```

Latent Class 1				
Thresholds				
Y6$1	3.133	0.537	5.831	0.000
Y7$1	36.977	0.000	999.000	999.000
Y8$1	2.717	0.388	7.004	0.000
Latent Class 2				
Thresholds				
Y6$1	16.781	0.000	999.000	999.000
Y7$1	18.488	0.000	999.000	999.000
Y8$1	1.477	0.758	1.947	0.051
Latent Class 3				
Thresholds				
Y6$1	−0.223	0.236	−0.945	0.345
Y7$1	**−0.680**	0.239	−2.842	0.004
Y8$1	−0.988	0.286	−3.450	0.001
Latent Class 4				
Thresholds				
Y6$1	0.175	0.416	0.420	0.675
Y7$1	−0.136	0.461	−0.296	0.768
Y8$1	0.806	0.282	2.856	0.004

不等值的模型自由度比未限定模型自由度增加了1个（130779−130778＝1），相应的G^2值增加了0.14（$p>0.05$），仅从拟合的角度考虑，不等值限制对模型拟合影响不大。

1.10　多组潜类别分析示例

多组LCA在M*plus* 中的设置类似于多种CFA的设置，具体来说，通过在VARIABLE下调用KNOWNCLASS命令来设置多组的信息，等同于多组CFA中的GROUPING命令。

也同多组CFA一样，可以通过限定相关参数来检验测量等值假设，框1-7列举了几个常见的多组LCA程序。框1-7中的第一部分设置gender变量为已知类别，即据此设置成多组分析。由于同时存在男女两个已知类别和四个潜类别，所以两个变量存在2（gender）×4（C）＝8个组合，17个条目在8组中分别有一个条件概率。如果不加任何参数限制，17个条目在8组中参数自由估计。如果通过框1-7第二部分的限制（MODEL：％overall％下设置相应参数等值），则设定8组参数等同。这样的限制模型最节俭，当然也最苛刻，更多时候，我们需要检验潜类别组在不同组之间（已知类别）的参数是否等值，所以需要释放部分参数等值的限定。具体通过在“Model C”下面设置各个类别条件概率自由估计，这样就检验了四个潜类别组的条件概率在男女两组间等值的假设是否成立（框1-7最后一部分）。

框1-7　多组CLCA的M*plus* 语句

```
TITLE: This is an example of a CLCA model with
   inequality restrictions
DATA: FILE IS PTSD__2. dat;
VARIABLE: NAMES ARE gender y0 y1-y17;
   USEVARIABLES ARE  y1-y17;
```

```
    CATEGORICAL＝y1-y17;
    CLASSES＝c（4）; gender（2）;
    KNOWNCLASS

＝gender（G1＝1 G2＝2）; ! 通过KNOWNCLASS
设置多组模型，非常类似于多组CFA时的Grouping命令;
  ANALYSIS: TYPE＝MIXTURE;
    STARTS＝20010;
    ALGORITHM＝INTEGRATION;
  OUTPUT: TECH11 TECH14;

  SAVEDATA: file＝ptsdclca-4-3. txt;
    save＝cprob;
  PLOT: type is plot3;
    series＝y1-y17（*

）;
! 设置所有类别下（即C×G＝4×2＝8种组合）的条件概率等值; MODEL:
%overall%
   [y1$1];
   [y2$1];
   [y3$1];
  …
   [y15$1];
   [y16$1];
   [y17$1];
```

```
! 设置C下的4个类别在潜类别G下的2个类别内的条件概率等值，当然也可以设置部分类别等值；
Model C：
%c#1%
  [y1$1]；
  [y2$1]；
  [y3$1]；
 …
  [y15$1]；
  [y16$1]；
  [y17$1]；
%c#2%
  [y1$1]；
  [y2$1]；
  [y3$1]；
 …
  [y14$1]；
  [y15$1]；
  [y16$1]；
  [y17$1]；
%c#3%
  [y1$1]；
  [y2$1]；
  [y3$1]；
 …
  [y15$1]；
  [y16$1]；
```

```
  [y17$1];
%c#4%
  [y1$1];
  [y2$1];
  [y3$1];
  …
  [y14$1];
  [y15$1];
  [y16$1];
  [y17$1];
```

注：限于篇幅，结果输出未呈现。

1.11　本章小结：潜类别分析作为混合模型的基本形式

探索性LCA作为经典的LCA已有几十年的历史。在经典LCA的基础上，新方法在不断发展。在LCA内部，逐渐发展出相应的多组潜类别模型和多潜变量的潜类别因子模型（LC factor models），以及多水平LCA。在外部，LCA连接着类别潜变量与连续潜变量模型。

第一，因子混合模型（Factor Mixture Model；e. g. ，Lubke＆Muthén，2005），其同时具有LCA和FA的特点。换句话说，在FMM中，同时用因子和潜类别两种潜变量对观测数据进行建模，通过同时抽取连续和间断的潜变量使观测指标达成局部独立性。

第二，将LCA与传统的回归模型相结合形成回归混合模型（Regression Mixture Model），即包含协变量的LCA和类别潜变量作为调节变量的回归混合模型。

第三，更一般的连续和类别潜变量结合的形式是结构方程混合模型（Structural Equation Mixture Modeling，SEMM；Bauer & Curran，2004；Muthén，2001）。因子混合模型和回归混合模型可以视作其特例。

LCA所处理的数据为横断面数据，当分析纵向数据时，称作潜在转换分析（Latent Transition Analysis，LTA；Collins & Lanza，2010），其目的在于刻画不同类别组个体随时间而改变类别属性的情况。其类型有结合传统的增长模型（Growth Model）的增长混合模型（Growth Mixture Modeling，GMM；Muthén & Muthén，2000；Muthén，2003）和潜类别增长模型（Latent Class Growth Analysis，LCGA；Nagin，1999）。

※推荐阅读※

Magidson＆Vermunt（2004）的文章简要介绍了LCM的原理和LCM的几种形式。Finch＆Bronk（2011）介绍了在M*plus* 上如何实现验证性LCA。中国台湾地区学者邱皓政于2008年在国内出版了第一本介绍LCA的专著《潜在类别模型的原理与技术》，该书从理论和操作两个方面对LCA的几种主要形式进行了介绍，同时使用LatentGOLD和M*plus* 进行分析，是非常不错的入门教材。Sterba（2013）最近的文章将LCM和

LPM放到混合模型中进行了介绍，是理解各种混合模型之间联系的优秀文章。

邱皓政（2008）．潜在类别模型的原理与技术．北京：教育科学出版社．

Finch，W．H．，＆Bronk，K．C．（2011）．Conducting Confirmatory Latent Class Analysis Using M*plus*．*Structural Equation Modeling*，18，132-151．

Magidson，J．，＆Vermunt，J．K．（2004）．Latent class models．In D．Kaplan．（Ed．），*The Sage Handbook of Quantitative Methodology for the Social Sciences*，（pp．175-198）．Thousand Oaks：Sage Publications．

Sterba，S．K．（2013）．Understanding Linkages Among Mixture Models．*Multivariate Behavioral Research*，48，775-815．

---

(1) 关于潜类别之间是否存在质的差异尚无定论。

(2) 对变量尺度的要求在M*plus* 框架内是相对宽松的，使用M*plus* 进行FA时对指标的要求也较宽松。

(3) 在M*plus* 中，两者的分析过程也非常类似，都是要逐步尝试不同的因子或类别个数，然后进行模型比较。

(4) Latent Class Factor Model可以直接译成潜类别因子模型，但为了与第3章的LCFM区别，在这里称作多个潜变量的潜类别模型。

(5) 本小结讨论的内容不仅适用于LCA，也适用于确定所有混合模型时的潜类别数目。

(6) 这里的条件概率为选项2上的概率。

2　回归混合模型

LCA和FA通常作为测量模型，因为两者都是处理潜变量和测量指标间关系的模型。同FA一样，LCA也可以进一步拓展，纳入协变量（预测变量和结局变量）。纳入协变量的FA称为SEM；纳入协变量的LCA称作回归混合模型（Regression Mixture Modeling，RMM；e. g. ，Clark＆Muthén，2009；Ding，2006）。例如，考察性别、种族等人口学变量对类别潜变量的影响。在某种意义上来说，RMM包含了SEM，或者说SEM只是RMM的一个特例。因为LCA模型纳入的变量可以是观测变量也可以是潜变量。如果纳入的是潜变量，RMM模型就变成了结构混合方程模型（Structure Mixture Equation Model，SMM），SEM变成了它的特例，即不存在类别潜变量的SMM。在本章中，我们先讨论协变量是观测变量的情况即RMM，然后再介绍SMM。

2.1　回归混合模型概述

模型包含的协变量通常存在两种类型[(1)]：预测变量（predictor variable）和结局变量（outcome variable or distal variable）。如图2-1所示：类别潜变量C 由测量指标U 测量；左图中预测变量x 指向类别潜变量C 的箭头表示协变量影响个体类别归属。例如，某研究者试图了解人口学变量对儿童行为问题潜在类别归属的影响，根据5个测量儿童行为问题的指标将450名儿童分成4个潜类别组（即潜类别变量“问题行为”有4个水平），然后构建人口学变量（性别、家庭经济地位和年龄等）对潜类别变量的回归模型。

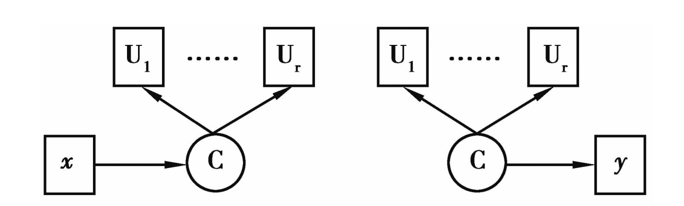


图2-1　回归混合模型示意图

在上图中，箭头的方向从潜类别变量C 指向结局变量y ，表示以类别属性（分类变量）预测结局变量。假设儿童潜类别归属可能会影响儿童学习成绩（由于成绩通常是连续变量，所以此时为线性回归），换句话说，不同问题行为类别的儿童学习成绩存在差异（根据类别潜变量将儿童分成4组然后做方差分析。当然，此时的方差分析和线性回归是等价的）。

在回归模型中，通常是根据因变量的类型选择对应的回归模型。上图中，类别变量C 通常有2个及以上水平，因此logistic回归和多项logistic回归是最常见的分析模型。下图的回归类型较为多样，主要取决于y 变量的类型，可能是线性回归也可能是其他形式的回归模型。

除了上述两种形式的RMM，还存在另外一种更一般的形式，如图2-2所示：类别潜变量C 有5个测量指标，自变量x 和因变量y 构成一个简单的回归模型，C 指向回归系数的虚线表示回归系数的大小在C 的不同水平是变化的。这里的类别潜变量C 可以理解为调节变量：变量x 和y 的关系在第三个变量的不同水平存在差异。

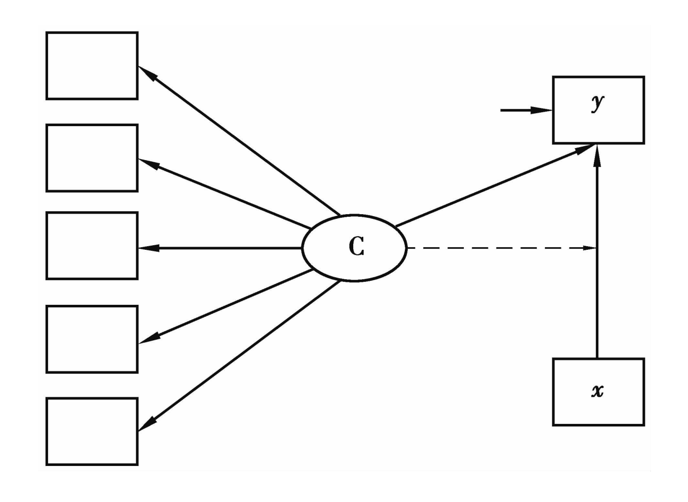


图2-2　潜类别变量作为调节变量的RMM

在RMM中，回归分析的逻辑很简单，类别潜变量作为分类变量纳入回归模型，但与传统回归模型又存在很大的差异。因为类别潜变量的各个类别水平是基于模型估计的，存在“测量误差”，而不能像观测类别变量那样不存在测量误差，可直接用于建模。如果潜在类别变量不存在分类误差，RMM就变成了传统的分组回归模型，但这种情况并不多见。在存在分类误差的前提下，协变量与类别潜变量之间的关系通常会被低估，误差越大低估越明显（Bolck，Croon＆Hagenaars，2004；Vermunt，2010）。

由于潜类别分组的不确定性（存在分类误差），近几年方法学者提出了不同的处理方法。总的来说，目前处理回归混合模型的方法尚处在发展阶段，新的方法不断地被提出和改进。本章主要介绍目前已知方法，重点介绍两种比较有效的方法：单步法和三步法。

根据模型中协变量的位置可以将回归混合模型分为3类：①只包含预测变量的RMM；②只包含结局变量的RMM；③潜类别变量作为调节变量的RMM。由于预测变量和结局变量有不同的类型（特别是结局变量的类型），所以方法亦不同。同时包含预测和结局变量的RMM，限于方法和软件的原因，目前尚无法实现。类别潜变量作为调节变量的RMM也较为常见，下面分别介绍3类不同的RMM。

2.2　包含预测变量的RMM

总的来说，带有预测变量的回归混合模型的建模策略可以大致分成两大类：单步法和分步法（三步法）。顾名思义，单步法在建模时

一步完成所有模型（测量和结构）参数估计；而分步法则采用逐步建模的步骤完成参数估计。由于单步法将看似不同的分析过程一步完成了，不像分步法过程清晰且易于理解。

（1）单步法

单步法在处理带有协变量的混合模型时，同时完成潜类别分组（测量模型部分）和协变量关系建模（结构模型部分）。如果协变量是预测变量，将其直接纳入模型进行分析，协变量与潜类别变量的关系在LCA分析中同步完成。

考虑协变量时的LCA表达式：

$$P(Y_i \mid Z_i) = \sum_{t=1}^{T} P(C = t \mid Z_i) \prod_{k=1}^{K} P(Y_{ik} \mid C = t) \tag{2.1}$$

$P（C=t / Z_i）$ 为考虑协变量Z时属于潜类别t的概率，该值可通过多项式logistic回归获得（Bakk＆Vermunt，2016）：

$$P(C = t \mid Z_i) = \frac{e^{\alpha_t + \beta_t Z_i}}{\sum_{s=1}^{T} e^{\alpha_s + \beta_s Z_i}} \tag{2.2}$$

上式中的 α_t 和 β_t 分别表示类别特定的截距和斜率。

如果协变量是结局变量（图2-1右图），只需将结局变量当作LCA的测量指标纳入模型（具体见后文）。然而单步法存在如下几点不足（Vermunt，2010）：

第一，当存在较多协变量时，单步法的实际操作性较差。在探索性研究中，由于缺少相关研究或理论预期，模型中常常包含多个预测变量。在单步法中，不同协变量的纳入和剔除都会影响测量模型（LCA）的结果，使得整个分析过程非常繁琐。

第二，模型建模困难。混合模型建模过程中最重要也是最复杂的问题是潜类别个数的确定（见本书第1章），包含协变量使得这一过程更加复杂。

第三，单步法在实践中不易被应用研究者理解和掌握。回归混合模型的逻辑顺序是先根据LCA将样本分组；接着以分组（潜）类别变量作为观测自变量或因变量进行回归分析，而在单步法中这些过程是一步完成的，理解和解释上较为抽象。

第四，包含协变量的LCA模型可能会违反混合模型的前提假设，例如协变量在类别内的方差相等或/和正态分布等（Bauer＆Curran，2003；Petras＆Masyn，2010）。

由于单步法的上述困难和不足，分析过程清晰的三步法受到方法学者和应用研究者的广泛关注（e．g．，Morin et al，2010；Chan＆Goldthorpe，2007）。

（2）简单三步法

按照大多数应用研究者的分析习惯，在进行混合模型分析时，通常根据多个测量指标采用LCA将样本分成不同的潜类别组（测量模型部分，见第1章）。然后将类别潜变量作为观测类别变量进行后续分析。

常见的后续分析有：比较变量在潜类别组上的差异（独立样本t检验或方差分析）；其他变量预测类别潜变量或类别潜变量预测其他变量。

三步法的分析过程如图2-3所示，分析由三步组成（三步法多数采用此流程）：

①进行常规的LCA模型估计，这一步只使用LCA的测量指标；

②接着在第一步的基础上根据后验概率获得个体的类别归属变量即潜类别分组变量；

③最后将潜类别分组变量作为观测变量（分类变量）连同协变量进行回归分析。

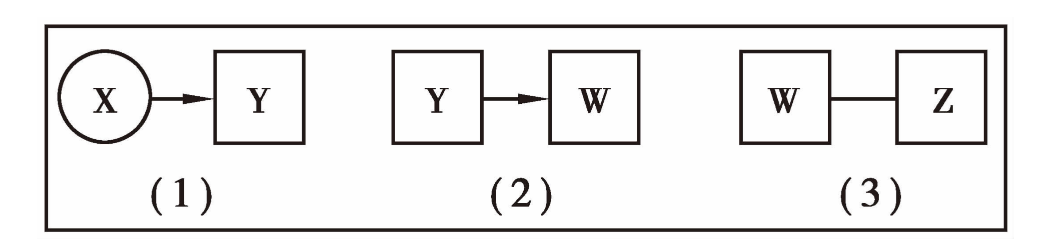


图2-3　简单三步法的分析流程

简单三步法也称作最可能类别回归法（Most Likely Class Regression；Clark＆Muthén，2009）[(2)]。这种方法符合应用研究者的分析习惯，在实践中广为使用。例如，在Chan和Goldthorpe（2007）研究社会分层与文化消费的文章中就采用了这种方法。他们首先根据6000多人在8个音乐消费指标上的反应数据，采用LCA将其分成4个潜类别，然后将这4个潜类别作为因变量进行多项式logit回归（Multinomial Logit Model）。

然而三步法也存在一些不足，通常是低估类别潜变量和协变量的关系，分类误差越大，系数低估越明显（Bolck et al，2004；Vermunt，2010）。回归混合模型分析的关键在于分类精确性（当然也是所有分类方法的关键）。分类精确性对于个体中心的方法来说可以理解为测量信度或测量误差问题。如果分类误差较大，把不属于某一类别的个体划分到该类别将会影响整个分析结果的准确性。针对简单三步法存在测量误差的问题，近年来研究者提出了一些校正方法（Bakk，Tekle & Vermunt，2013；Lanza，Tan & Bray，2013；Vermunt，2010），下面将逐一详细介绍。

（3）概率回归法和加权概率回归法

这两种方法的分析过程与简单三步法类似，也是分成三步。具体来说，第一步依据观测指标将个体分类，即执行LCA分析。第二步将个体的后验概率进行转换后再做回归分析：①概率回归法将后验概率进行对数转换，转换后的数值作为结果进行回归分析；②加权概率回归法则根据后验分类结果直接与协变量进行回归但采用后验概率进行加权。两种方法都考虑了分类的不确定性，与简单三步法相比回归系数的结果较为准确，但由于后验概率的估计本身也是存在误差的，所以回归系数的显著性检验存在结论错误的可能（Clark & Muthén，2009）。框2-1和框2-2分别给出了两种方法的M*plus* 例句（Clark & Muthén，2009）。

框2-1　概率回归法和概率加权回归法的第一步估计的M*plus* 语句

免0无法计算。

```
Title: the step1 for probabilities regression model
Data: File is sim1. dat;
Variable: Names are u1-u5 x;
Usevariables＝u1-u5;
Categorical＝u1-u5;
Classes＝c1（2）;
Auxiliary＝x;   ! 辅助变量将被保存到结果文件用于后续分析
Analysis: Type＝Mixture;
    Starts＝0;
Model: %Overall%
Output: Tech11;
Savedata: File is simpp1. dat; ! 保存结果的文件名
Save＝Cprob; ! 保存后验概率
```

框2-2　概率回归和概率加权回归法的*Mplus* 语句

```
Title: probabilities regression model
Data: File is simpplist. dat;
Variable: Names＝u1-u5 x p1 p2 n; ! p1 p2为两个类别变量的后验概率，如果第一步估计n个类别，这里的p就有n个。
Usevariables＝x logit;
Define: ! 通过定义命令将概率转换成logit形式
If（p1 lt . 00001）then p1＝. 00001; ! 将非常小的值定义为. 00001避免0无法计算。
```

```
If（p1 gt . 9999）then p1＝. 9999；！将非常大的值定义为. 9999避免1无法计算。
logit＝log（p1／（1-p1））；！转换成对数形式
Model: logit on x；！定义回归模型


Title: probabilities as weights in a regression.
Data: File is simpplist. dat;
Variable: Names＝u1-u5 x p1 p2;
  Usevariables are u1-u5 x p1 p2;
  Categorical＝u1-u5;
  Classes＝c1（2）；
  Training＝p1 p2（probabilities）；！通过Training命令定义p1 p2为权重
Analysis: Type＝Mixture;
Model:
  %Overall%
  c1 on x;
```

（4）虚拟类别法

LCA根据一次分析的后验概率将个体分组，这种做法存在抽样误差的问题[(3)]。虚拟类别法（Pseudoclass Method，PC法）采用类似缺失值分析时使用的多重插补法，从个体的后验概率分布中随机抽取若干个（通常20次）可能的后验概率值[(4)]，根据每次的概率值将个体

分配到不同的类别，然后平均若干次的结果作为最终的分类结果（Wang，Brown＆Bandeen-Roche，2005）。

Clark＆Muthén（2009）的模拟发现，当分类精确性较高时（entropy＞0.8），该方法表现较好；然而在最近的模拟研究中发现，与稳健三步法和单步法相比，虚拟类别法在同等条件下表现最差（Asparouhov＆Muthén，2014），在实际应用中并不被推荐使用。框2-3给出了虚拟类别法的M*plus* 例句。

框2-3　虚拟类别法在M*plus* 中的设定

```
TITLE: Input file for pseudoclass method
DATA: file＝man3step. dat;
VARIABLE: Names＝u1-u5 y x;
USEVARIABLE＝u1-u5;
Categorical＝u1-u5;
Classes＝c（3）;
Auxiliary＝x（r）; ! 辅助变量后加（r）采用的是PC方法
ANALYSIS: Type＝Mixture;
SAVEDATA: FILE＝man3step2. dat; SAVE＝CPROB;
```

（5）稳健三步法或MML法

稳健三步法由Vermunt（2010）在Bolck，Croon & Hagenaars（2004）研究的基础上提出。由于同时采用莫代尔法分配法

和极大似然估计，因此又称为莫代尔极大似然估计法（Modal ML）。Asparouhov 和 Muthén （ 2014 ） 将 其 称 作 三 步 法 （ 3-steps approach），为了区分简单三步法，我们在这里将其称作稳健三步法。分析步骤同简单三步法，区别在于第二步考虑了分类误差，而简单三步法并未处理分类误差。稳健三步法的具体分析步骤如图2-4所示。

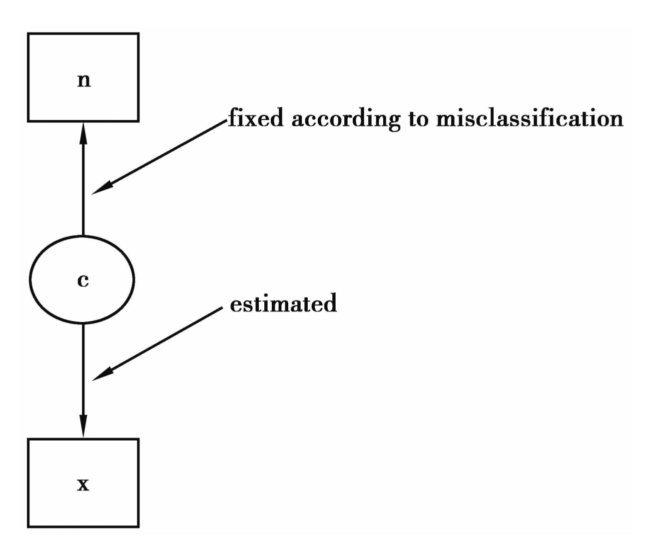


图2-4　稳健三步法分析流程图

稳健三步法最大的特点是在第二步考虑了分类误差或不确定性。假设W是基于模型估计的类别潜变量，与实际的类别潜变量C 并不完全一致（完全一致时不存在分类误差），因此存在如下2个分类不确定率：

$$p_{c_1,c_2} = P(C = c_2 \mid N = c_1) = \frac{1}{N_{c_1}} \sum_{N_i = c_1} P(C_i = c_2 \mid U_i) \tag{2.3}$$

上式中，C 为类别潜变量，N 为根据后验分布概率将个体划分到不同潜类别组的变量（M*plus* 分析无条件LCA模型时保存后验概率后结果文

件的最后一列），U 为观测指标。N_{c_1} 是根据N 将个体划分到C_1 类别的数量。

在M*plus* 的新近版本中（7.2之后的版本），p_{c_1,c_2} 的值可以在结果输出部分获得。随后可以计算“分类错误率”：$P(N=c_1 \mid C=c_2)$ 即实际属于C_2 类别但在LCA中根据后验概率却被归入C_1 的概率：

$$q_{c_2,c_1}=P(N=c_1 \mid C=c_2)=\frac{p_{c_1,c_2}N_{c_1}}{\sum_c p_{c,c_2}N_c} \tag{2.4}$$

N_c 是根据N 将个体分配到C 的数量。稳健三步法使用$\log(q_{c_1,c_2}/q_{k,c_2})$ 作为N 估计C的权重，该值可在M*plus* 结果报告中直接获得。

在M*plus* 中，稳健三步法有两种实现形式：自动和手动。自动形式只需采用AUXILIARY的R3STEP选项，软件自动完成上述3步分析。手动形式需要分别执行两步分析。第一步，单独执行LCA分析，获得分类错误率的对数形式。第二步，在这一步分析中，将第一步保留的分组变量N的均值固定为分类错误率的对数值。框2-4和框2-5分别给出了稳健三步法自动和手动两种实现形式的例句。

框2-4　稳健三步法自动实现的M*plus* 语句

```
DATA：FILE＝3step. dat;
VARIABLE：NAMES＝u1-u5 x;
CATEGORICAL＝u1-u5;
```

```
CLASSES＝c（3）；
AUXILIARY＝x（R3STEP）；！通过设置（R3STEP），X作为预测变量采用稳健三步法估计。
ANALYSIS：TYPE＝MIXTURE；
MODEL：！no model is needed，LCA is default
```

框2-5　手动执行稳健三步法的M*plus* 语句

```
Title：Input for step 1 in the manual 3-step estimation
DATA：file＝man3step. dat；
VARIABLE：Names are u1-u10 y x；
    usevariable＝u1-u10；
    Categorical＝u1-u10；
    Classes＝c（3）；
    auxiliary＝y x；
ANALYSIS：Type＝Mixture；
SAVEDATA：FILE＝man3step2. dat；SAVE＝CPROB；
Title：Input for step 3 in the manual 3-step estimation
DATA：file＝man3step2. dat；
VARIABLE：Names are u1-u10 y x p1-p3 n；
    usevar＝y x n；
    classes＝c（3）；
    nominal＝n；
ANALYSIS：Type＝Mixture；
    starts＝20050；
```

```
MODEL：%overall%
    C on X；
    %C#1%
     [N#1@1.901]；! 数值1.901根据第一步的结果获得
     [N#2@-0.990]；
    %C#2%
     [N#1@-0.486]；
     [N#2@1.936]；
    %C#3%
     [N#1@-2.100]；
     [N#2@-2.147]；
```

（6）修正的BCH法

BCH法最早由Bolck，Croon＆Hagenaars（2004）提出，用于处理包含分类预测变量的LCA。该方法与稳健三步法逻辑类似，区别在于稳健三步法的第三步的估计方程采用极大似然估计，而BCH将其转换成加权方差分析，分类误差作为权重。

与稳健三步法相比，BCH法的一个突出优点是不会改变潜类别的顺序。潜类别顺序的改变是极大似然估计的一个“副产品”。在上一章我们提到，由于ML估计常常得到局部最大化解而非整体最大化解，所以混合模型估计通常设置多个起始值，而起始值通常由软件随机生成，所以每次分析的起始值不同得到的潜类别结果可能不同，潜类别的顺序也可能不同。尽管使用相同的数据和指标，所得到的拟合结果

和类别数目也相同，但类别潜变量水平的顺序可能不同（第一个类别变成第二个类别），因此给潜类别分析带来很大的麻烦。

在稳健三步法分析中，M*plus* 自动监测顺序改变问题，一旦发生顺序改变，M*plus* 将不报告结果（Asparouhov＆Muthen，2015；webnote＃21）。BCH法的不足在于，当类别距离很小及样本量小时，类别内的误差方差可能是负值。此时如果把类别内方差固定相等，也可以获得正确的类别组内结局变量的均值（Bakk＆Vermunt，2014）。

框2-6　BCH法执行带有预测变量RMM的M*plus* 语句

```
Title：step 1 in the BCH estimation for RMMwith predictor；
DATA：file＝manBCH. dat；
VARIABLE：Names＝U1-U10 Y X；
    Categorical＝U1-U10；
    Classes＝C（3）；
    Usevariable＝U1-U10；
    Auxiliary＝Y X；
ANALYSIS：Type＝Mixture；
SAVEDATA：File＝BCH2. dat；
    Save＝bchweights；
Title：step 2 in the BCH estimation for RMMwith predictor；
Data：file＝BCH2. dat；
Variable：Names＝U1-U10 Y X W1-W3 MLC；
    usevariable＝X W1-W3；
    Classes＝C（3）；
```

```
  Training＝W1-W3（BCH）；
Analysis：Type＝Mixture；
  Starts＝0；
  Estimator＝MLR；
Model：％overall％
  C on X；
```

（7）小结

就目前的模拟研究结果来看，稳健三步法和单步法是处理含有预测变量RMM最好的方法。根据通常的潜类别建模流程，首先确定群体分类，然后再在此基础上做进一步分析。稳健三步法的分析过程清晰明确，因符合广大应用研究者的分析习惯而容易被接受。

2.3　包含结局变量的RMM

总的来说，包含结局变量的RMM比包含预测变量的RMM要复杂一些，因为在后者的建模过程中类别潜变量作为因变量只存在一种形式——logistic或多项式logistic回归。但在包含结局变量的RMM中，结局变量存在两种形式：连续和类别变量。下面分别介绍两种不同形式结局变量的RMM分析方法。

2.3.1　结局变量是连续变量

（1）单步法

结局变量是连续变量时，可以将结局变量当作LCA模型的指标，同时完成模型估计。当局部独立性满足时，LCA表达式如下：

$$P(Y_i) = \sum_{t=1}^{T} P(C=t) \prod_{k=1}^{K} P(Y_{ik} \mid C=t) \qquad (2.5)$$

当纳入连续的协变量Z后，上式改写为联合的形式：

$$P(Y_i \mid Z_i) = \sum_{t=1}^{T} P(C=t) \prod_{k=1}^{K} P(Y_{ik} \mid C=t) f(Z_i \mid C=t) \qquad (2.6)$$

$f（Z_i \ /X=t）$ 为协变量Z 在特定类别内的分布，连续变量时为正态分布，如果存在多个连续变量则为多元正态分布。

单步法需要满足重要的前提：连续结局变量在各类别内正态分布。如果正态假设不成立则会改变测量模型的结构及意义，例如高估类别数（Bauer＆Curran，2003）。另外，如果存在多个连续结局变量则更加复杂。假如采用每次只纳入一个结局变量的建模策略，则会产生LCA模型混淆的问题：纳入不同结局变量间的LCA模型是不同的。

（2）LTB法

Lanza，Tan＆Bray（2013）最近提出了一种新的方法可以避免单步法违反假设时结果不准确的问题，因为这种方法并没有特定的分布

假设。在LTB法中，首先将结局变量Z作为协变量纳入LCA分析（过程同包含预测变量的单步法）。

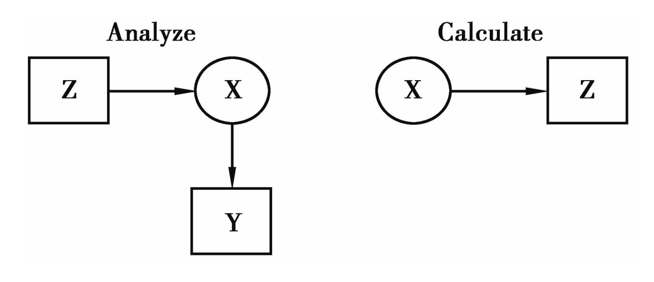


图2-5 LTB法示意图

第二步计算结局变量在每个类别内的均值（连续变量）或概率（类别变量）[5]：

$$\mu_t = \int_Z Zf(Z \mid C = t) \qquad (2.7)$$

协变量Z在特定类别内的分布$f\,(Z\,/\,X=t)$ 可通过贝叶斯定理获得：

$$f(Z \mid C = t) = \frac{f(Z)P(C = t \mid Z)}{P(C = t)} \qquad (2.8)$$

其中的$P\,(X=t\,/\,Z)$ 和$P\,(X=t)$ 是条件概率和类别概率，由第一步获得。式中$f\,(Z)$ 未知，Asparouhov和Muthén（2014）建议使用Z的实证分布代替：

$$\mu_t = \sum_{i=1}^{N} Z_i \frac{P(C = t \mid Z_i)}{NP(C = t)} \qquad (2.9)$$

Lanza，Tan & Bray（2013）并没有给出μ_t的标准误公式，Asparouhov和Muthén（2014）建议使用类别特定的方差的均方根除以

类别特定的样本量获得，但模拟研究发现这种做法会低估标准误（Bakk＆Vermunt，2016）。随后，Bakk，Oberski＆Vermunt（2016）提出了Jackknife和Bootstrap再抽样的标准误。

当连续结局变量的方差在不同类别内相等时即同方差（homoskedastic），LTB法的估计结果是无偏的，此时结局变量与潜类别变量之间呈linear-logistic关系。如果同方差不成立即异方差（heteroskedastic）时，LTB法估计类别特定的均值存在偏差（Bakk＆Vermunt，2016）。另外，LTB方法处理多个连续结局变量时存在困难，如果采用分别建模的方式将面临与单步法同样的困境。

框2-7　LTB法分析带有连续结局变量RMM的M*plus* 语句

```
Title： the LTB estimation for RMMwith categorical distal variable;
Data: File＝8-data. dat;
Variable: Names are U1-U8 Y;
    Categorical＝U1-U8;
    Classes＝C（4）;
Auxiliary＝Y（DCONTINUOUS）;
Analysis: Type＝Mixture;
```

（3）修正的LTB法

针对LTB法的不足，Bakk，Oberski＆Vermunt（2016）结合稳健三步法的分析思想对LTB法进行了修正，并将其分成三步实现，因此这种方法与稳健三步法分析过程非常相似。首先，使用测量指标建立LCA，同时根据后验概率将个体分到不同的潜类别组N。第二步，考虑分类误差的前提下通过N估计潜类别变量C，同时将结局变量Z作为协变量纳入分析（稳健三步法并未纳入协变量），见公式2.10。

$$P(N_i = s \mid Z_i) = \sum^{T} P(C = t \mid Z_i) P(N_i = s \mid C = t) \quad (2.10)$$

上式中的$P(N_i = s / C = t)$被固定为上一步估计的N。

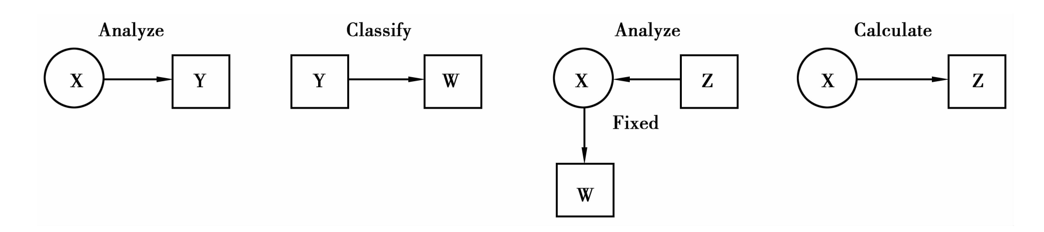


图2-6　修正的LTB法示意图

当连续结局变量的方差在不同类别内不相等时（类别内异方差），LTB法的估计结果是有偏的。针对此问题，Bakk，Oberski＆Vermunt（2016）提出在多项式logistic回归模型中加入二次项（公式2.11）来解决估计偏差的问题。

$$P(C = t \mid Z_i) = \frac{\exp(\alpha_t + \beta_t Z_i + \gamma_t Z_i^2)}{\Sigma_{t'=1}^{T} \exp(\alpha_{t'} + \beta_{t'} Z_i + \gamma_t Z_i^2)} \quad (2.11)$$

修正的三步LTB法才刚刚提出，还需要更多方法学研究，当然尚未有实际应用的文章发表。

（4）修正BCH法

如前所述，BCH法提出之初仅用于分析包含分类预测变量的LCA，后来Vermunt（2010）对其进行了修正，使其可以处理各种类型的变量。在M*plus* 里，使用BCH分析包含结局变量RMM时非常方便，只需一步即可实现，例句见框2-8。

框2-8　BCH法分析带有结局变量RMM的M*plus* 语句

```
Title: the BCH estimation for RMMwith distal variable;
Data: File=8-data. dat;
Variable: Names are U1-U8 Y;
    Categorical=U1-U8;
    Classes=C（4）;
Auxiliary=Y（BCH）;
Analysis: Type=Mixture;
```

（5）稳健三步法

稳健三步法也可以用于处理结局变量是连续变量的RMM。包含连续结局变量的LCA模型表达式变为：

$$P(N=s \mid Z_i)=\sum_{t=1}^{T} P(C=t) f(Z_i \mid C=t) P(N=s \mid C=t) \quad (2.12)$$

$P(N=s|C=t)$ 被固定为第二步估计的分类精确性参数，$f(Z_i|C=t)$ 通常服从正态分布。如前所述，结局变量是连续变量的RMM的目的在于估计结局变量在潜类别不同水平上的均值差异，但结局变量的方差在不同类别组内可能相等也可能存在差异（类似方差分析时的组内方差同质假设）。针对方差的不同情况，稳健三步法有两种不同的变式：类别组内方差同质和类别组内方差异质。

框2-9　稳健三步法（等方差和不等方差）处理连续结局变量的M*plus* 语句

```
TITLE: 3-step method done automatically using DU3STEP:
VARIABLE: NAMES＝u1-u5 x;
CATEGORICAL＝u1-u5;
CLASSES＝c（3）;
AUXILIARY＝x（DU3STEP）;
DATA: FILE＝3 step. dat;
ANALYSIS: TYPE＝MIXTURE;

TITLE: 3-step method done automatically using DE3STEP:
VARIABLE: NAMES＝u1-u5 x;
CATEGORICAL＝u1-u5;
CLASSES＝c（3）;
AUXILIARY＝x（DE3STEP）;
DATA: FILE＝3 step. dat;
ANALYSIS: TYPE＝MIXTURE;
```

（6）几种方法的比较

模拟研究发现（Bakk et al，2013；Lanza et al，2013），当满足假设条件时[(6)]，稳健三步法，BCH和LTB均可以得到无偏的参数估计结果（即类别特定的结局变量的均值）。然而，当条件不成立时（非正态、方差不同质），稳健三步法和LTB表现较差，而BCH法则表现得很稳健（Bakk＆Vermunt，2016）。Asparouhov和Muthén（2015；webnote 21）通过模拟进一步比较了稳健三步法的两种变式（即类别等方差和类别不等方差，分别对应M*plus* 中的DE3STE*P* 和DU3STE*P* ），LTB法，单步法，PC法和BCH法在连续结局变量非正态（双峰分布）时的表现，结果进一步证实了BCH的稳健性（其他方法表现均不佳）。尽管如此，当类别距离或分类精确性较小时（比如entrop*y* ＝0.5），BCH也会低估标准误。他们的结果还发现，当组内方差同质性不成立时，方差不等的稳健三步法（DU3STE*P* ）和BCH法表现最佳，且前者更优。

2.3.2　结局变量是类别变量

LTB法在处理分类结局变量时表现较好，不会像分析连续结局变量时出现违反正态和方差同质假设后的估计偏差问题。在Asparouhov＆Muthén（2014）的模拟研究中，检验了3个样本量（*N* ＝200,500和2000）和2种分类精确性（entropy＝0.5和0.65）下LTB的表现，结果发现仅在*N* ＝200和entropy＝0.5时才会出现明显的偏差。

框2-10　LTB法分析带有类别结局变量RMM的M*plus* 语句

```
Title： the LTB estimation for RMMwith categorical distal
variable;
Data: File＝8-data. dat;
Variable: Names are U1-U8 Y;
    Categorical＝U1-U8;
    Classes＝C（4）;
Auxiliary＝Y（DCATEGORICAL）;
Analysis: Type＝Mixture;
```

2.4　潜类别变量作为调节变量的RMM

如图2-2所示的潜类别变量作为调节变量的RMM，这种类型的模型常称为混合回归模型（Mixture Regression Model，MRM），在实际研究中较为常见，而且更加符合RMM的本意，本章介绍的其他模型只能称为带有协变量的LCA。RMM更一般的形式是混合结构方程模型。

通常这类模型包含两部分：LCA和回归模型。前者是一个独立的测量模型，类别潜变量C 有5个测量指标；后者是常见的回归模型。同多组潜变量模型（分组变量为观测变量）一样，潜类别变量作为调节变量的RMM可以看作调节模型：类别潜变量就像分组变量一样调节不同类别的回归系数。采用这种方法的应用研究如Van Horn等（2009）在一个大样本N ＝6305中发现家庭资源对学业成绩的影响存在群体异质

性，换句话说，家庭资源对学业成绩的影响在不同类别组内存在差异。

传统的线性回归模型可以通过下式表达：

$$y = \alpha + \beta x + \varepsilon$$

同FMM一样，RMM可以看作传统回归模型与潜类别模型结合的产物（Lazarsfeld＆Henry，1968；McCutcheon，1987）。由此，RMM的表达式变成了如下形式：

$$y = \alpha_c + \beta_c x + \varepsilon_c$$

与传统的回归模型不同，截距和斜率系数多了下标c，表达截距和斜率跨类别变化。

这类模型的估计可以采用三步法实现。首先，估计LCA模型，将样本分组；其次，由于LCA分组存在分类错误，因此在随后的分析中要将分类错误考虑进来。最后，进行分组回归分析。回归混合模型的三步法目前无法在M*plus* 中同时建模，只能手动分步计算。

（1）单步法

单步法处理潜类别变量作为调节变量的RMM示例见框2-11。

框2-11　单步法处理潜类别变量作为调节变量的RMM

```
Title: one step method for C as moderator
DATA: file=step. dat;
VARIABLE: Names are u1-u10 y x;
    Categorical=u1-u10;
    Classes=c (3) ;
    Usevariable are u1-u10 y x;
Analysis: Type=Mixture;
    starts=20050;
Model:
    %Overall%
    Y on x;
    %c#1%
     [u1$1-u10$1*

-1] ;
    %c#2%
     [u1$1-u10$1*

1] ;
    %c#3%
     [u1$1-u5$1*

1] ;
     [u6$1-u10$1*
```

```
-1]；
SAVEDATA：FILE＝onestep. dat；SAVE＝CPROB；
```

（2）稳健三步法

在M*plus* 里采用稳健三步法处理潜类别变量作为调节变量的RMM需要两步完成。第一步分析LCA并保存后验分类概率；第二步，考虑分类误差进行结构模型分析，框2-12给出了两步的示例。

框2-12　类别潜变量作为调节变量的RMM的稳健三步法语句

```
Title：Input for step 1 in the manual 3-step estimation
DATA：file＝man3step. dat；
VARIABLE：Names are u1-u10 y x；
    Categorical＝u1-u10；
    Classes＝c（3）；
    usevariable are u1-u10；
    auxiliary＝y x；
    Analysis：Type＝Mixture；starts＝0；
    Model：
    %Overall%
    %c#1%
     [u1$1-u10$1*
```

```
-1];
    %c#2%
     [u1$1-u10$1*


1];
    %c#3%
     [u1$1-u5$1*


1]; [u6$1-u10$1*


-1];
  SAVEDATA: FILE=man3step2. dat; SAVE=CPROB;


Title: Input for step 3 in the manual 3-step estimation
    data: File=man3step2. dat;
    variable: Names are u1-u10 y x p1-p3 n;
    usevariable are y x n;
    classes=c(3);
    nominal=n;
    Analysis: Type=Mixture; starts=0;
    Model:
    %overall%
    Y on X;
    %C#1%
     [N#1@1.901]; [N#2@-0.990];
    Y on X; Y;
```

```
%C#2%
 [N#1@-0.486]; [N#2@1.936];
Y on X; Y;
%C#3%
 [N#1@-2.100]; [N#2@-2.147];
Y on X; Y;
```

（3）BCH法

与稳健三步法类似，BCH法处理潜类别变量作为调节变量的RMM也需要两步完成。第一步分析单独的LCA，保留后验概率作为第二步分析的权重，具体使用SAVEDATA里的Save＝bchweights。第二步使用Training语句将上一步保留的后验分类概率作为BCH法的权重，见框2-13的示例。

框2-13　BCH法执行类别潜变量作为调节变量的RMM的M*plus* 语句

```
Title: step 1 in the BCH estimation for C as moderator;
DATA: file=BCH. dat;
VARIABLE: Names=U1-U10 Y X;
    Categorical=U1-U10;
    Classes=C（3）;
    usevariable=U1-U10;
    Auxiliary=Y X;
```

```
ANALYSIS: Type=Mixture;
SAVEDATA: File=BCH2. dat;
    Save=bchweights;
Title: step 2 in the BCH estimation for C as moderator;
Data: file=BCH2. dat;
Variable: Names=U1-U10 Y X W1-W3 MLC;
    usevariable=Y X W1-W3;
    Classes=C (3) ;
  Training=W1-W3 (BCH) ;
Analysis: Type=Mixture;
  Starts=0;
  Estimator=MLR;
Model:
  %overall%
  Y on X;
  %C#1%
  Y on X;
  %C#2%
  Y on X;
  %C#3%
  Y on X;
```

2.5 混合结构方程模型

上述介绍的RMM有一个共同点，即类别潜变量都有自己的测量指标，协变量和结局变量通常不参与类别潜变量的计算（单步法除外）。而在混合结构方程模型中（如图2-7），类别潜变量没有自己的测量指标，更确切地说，是根据模型中所有指标的数据进行潜在分组。

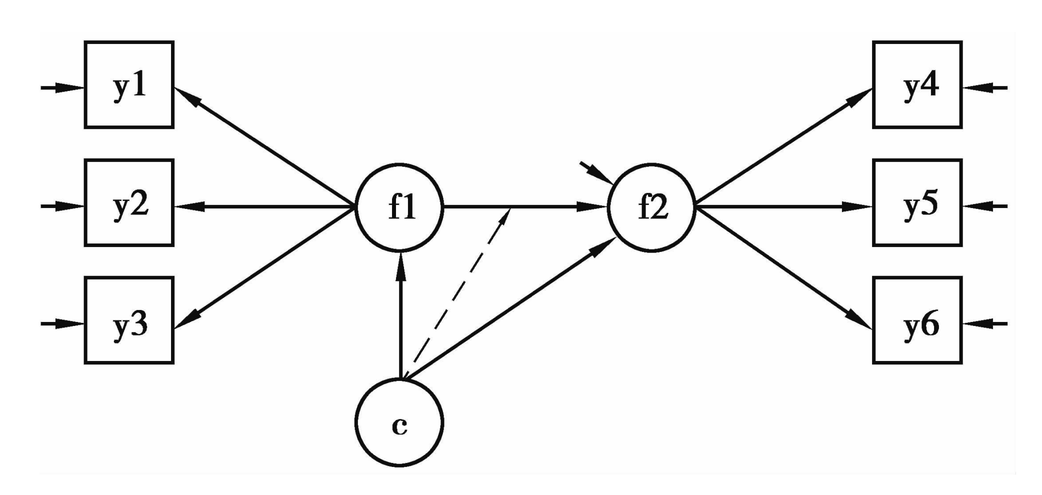


图2-7　混合结构方程模型示意图（来源：M*plus* 手册）

框2-14　混合结构方程模型的M*plus* 语句

```
TITLE： this is an example of structural equation mixture
modeling
DATA：FILE IS ex7.20. dat；
VARIABLE：NAMES ARE y1-y6；
  CLASSES＝c（2）；
ANALYSIS：TYPE＝MIXTURE；
MODEL：％OVERALL％
  F1 BY y1-y3；    F2 BY y4-y6；
  F2 ON f1；
  ％c＃1％
```

```
  [F1□1 F2];
 F2 ON F1;
OUTPUT: TECH1 TECH8;
```

来源：M*plus* 手册

2.6　实例分析

实例数据来自中国人民大学2010—2011年执行的北京市城镇老年人（60～95岁）焦虑症状调查。有效样本为1292个。本例中使用了其中的简版老年抑郁量表（GDS-15）总分（gds）、生活自理状况［共16个题项（C2A—C2Q），选项编码为：1—不费力；2—有些困难；3—做不了］、年龄（连续变量）、“觉得自己现在老吗”（ifold）（二分变量）等题目。

（1）潜类分析

首先，使用老年人生活自理状况量表的15个条目进行潜类别分析。分别拟合2～4个类别。通过模型比较后选择两个类别模型为最优模型（代码见框2-15）。

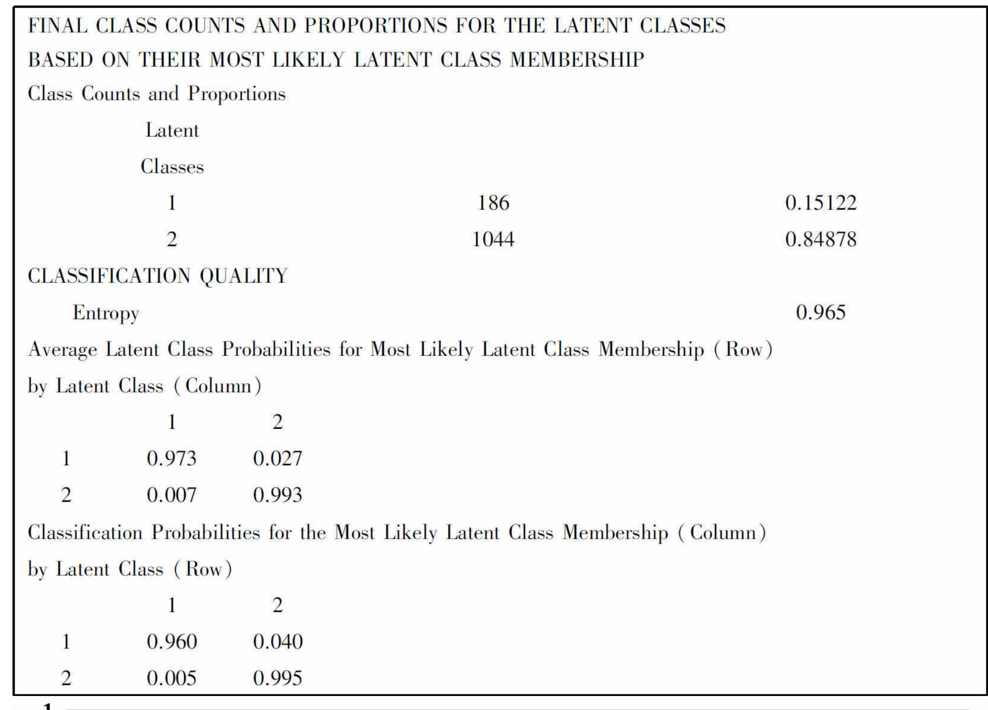
框2-15　示例潜类别分析M*plus* 语句

```
Title: Lantent Class Analysis
Data: File is older__survey. dat;
Variable: Names＝C2A C2B C2C C2D C2E C2F C2G C2H C2I C2J C2K
C2L C2MC2N C2P C2Q ifold age gds;
  USEVARIABLES＝C2A-C2Q;
  MISSING are all（-9999）;
  CATEGORICAL＝C2A-C2Q;
  CLASSES＝C（2）;
Analysis:
  TYPE＝MIXTURE;
  Starts＝50 3;
  PROCESSORS＝4;   ！根据电脑情况指定
PLOT:
  TYPE＝PLOT3;
  SERIES＝C2A-C2Q（*

）;
```

框2-16为潜类别分析输出结果。从结果来看，分为两类的结果拟合情况良好（相对于3类或4类），Entropy＝0.965。两类个体从实际意义来看，可以命名为“不能自理类”和“能够自理类”，分别占比15.1％和84.9％。图2-8呈现了两个类别的条件概率。

框2-16　潜类别分析输出结果（部分）



FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES
BASED ON THEIR MOST LIKELY LATENT CLASS MEMBERSHIP
Class Counts and Proportions

Latent Classes		
1	186	0.15122
2	1044	0.84878

CLASSIFICATION QUALITY

Entropy	0.965

Average Latent Class Probabilities for Most Likely Latent Class Membership (Row)
by Latent Class (Column)

	1	2
1	0.973	0.027
2	0.007	0.993

Classification Probabilities for the Most Likely Latent Class Membership (Column)
by Latent Class (Row)

	1	2
1	0.960	0.040
2	0.005	0.995

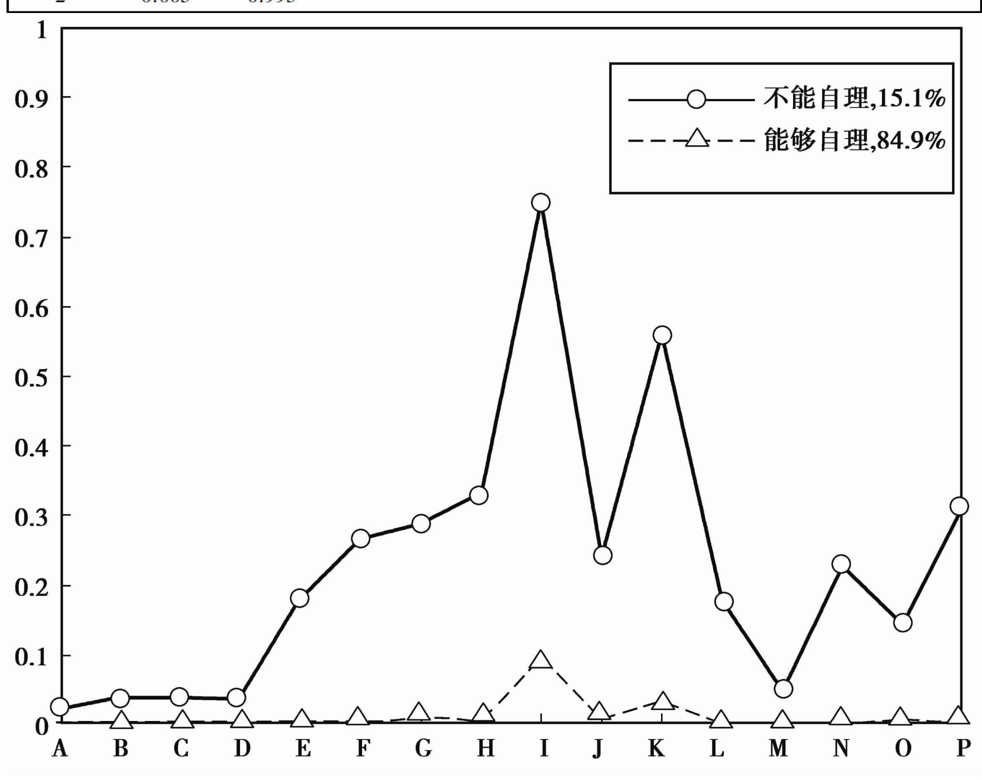


图2-8　两类别在选项3上的条件概率

（2）加入预测变量的回归混合模型

在保留的两个类别模型基础上加入连续预测变量（年龄），预测潜类别变量，采用R3STE*P* 法，相应的M*plus* 语句见框2-17。

框2-17　加入预测变量回归混合模型的M*plus* 语句

```
Title: Regression Mixture Modeling with Predictive Variable
Data: File is older__survey. dat;
Variable:  Names＝C2A C2B C2C C2D C2E C2F C2G C2H C2I C2J C2K
C2L C2M C2N C2P C2Q ifold age gds;
  USEVARIABLES＝C2A-C2Q;
  MISSING are all（-9999）;
  CATEGORICAL＝C2A-C2Q;
  CLASSES＝C（2）;
  AUXILIARY＝age（R3STEP）;
Analysis:
  TYPE＝MIXTURE;
  PROCESSORS＝4;
PLOT: TYPE＝PLOT3;
SERIES＝C2A-C2Q（*

）;
```

如前所述，因变量为分类潜变量，这里的回归方程为多项式logistic回归。软件默认第2个类别组为参照组（reference group）。从框2-18的结果来看，年龄对第一个类别的回归系数为

0.153，SE＝0.014，p ＜0.001。结果表明，年龄有助于预测老人所属的类别组。相对于第二类别组（能自理组），年龄大一岁属于第一类别的发生比要高出16.5个百分点（（$e^{0.153}$ －1）×100＝16.5）。

框2-18　加入预测变量回归混合模型输出结果（部分）

```
TESTS OF CATEGORICAL LATENT VARIABLE MULTINOMIAL LOGISTIC REGRESSIONS USING
THE 3-STEP PROCEDURE
                                                          Two-Tailed
                          Estimate      S.E.    Est./S.E.     P-Value
     C#1      ON
        AGE                 0.153       0.014      11.219       0.000
      Intercepts
        C#1               -12.935       1.031     -12.541       0.000
Parameterization using Reference Class 1
     C#2      ON
        AGE                -0.153       0.014     -11.219       0.000
Intercepts
        C#2                12.935       1.031      12.541       0.000
```

（3）加入分类结局变量的回归混合模型

同样地，在保留两个类别模型基础上加入自我感觉“是否老了”作为结局变量。该变量有2个选项，所以采用DCAT法，语句见框2-19。

框2-19　加入分类结局变量回归混合模型的*Mplus* 语句

```
Title： Regression Mixture Modeling with categorical outcome
variable
Data: File is older__survey. dat;
Variable: Names＝C2A C2B C2C C2D C2E C2F C2G C2H C2I C2J C2K
```

```
C2L C2MC2N C2P C2Q ifold age gds;
  USEVARIABLES＝C2A-C2Q;
  MISSING are all（-9999）;
  CATEGORICAL＝C2A-C2Q;
  CLASSES＝C（2）;
  AUXILIARY＝ifold  （DCAT）;
Analysis:
  TYPE＝MIXTURE;
  PROCESSORS＝4;
  LRTSTARTS＝2 1 80 16;
PLOT:
  TYPE＝PLOT3;
  SERIES＝C2A-C2Q（*

）;
```

对于分类结局变量的回归混合模型，M*plus* 输出结果中会出现跨类概率相等性检验，其中会输出相关的概率值。从框2-20的结果来看，相比于生活自理类别组，生活不能自理类别的老人其“老人身份认同”的程度要高。具体结果是，生活不自理类别组选择“觉得自己老了”的概率是0.735，“觉得自己未老”的概率是0.265；而生活能自理组对应的选择是0.435和0.565。

框2-20　加入分类结局变量回归混合模型输出结果（部分）

EQUALITY TESTS OF MEANS/PROBABILITIES ACROSS CLASSES						
IFOLD						
	Prob	S.E.	Odds Ratio	S.E.	2.5% C.I.	97.5% C.I.
Class 1						
Category 1	0.265	0.033	1.000	0.000	1.000	1.000
Category 2	0.735	0.033	2.133	0.389	1.492	3.049
Class 2						
Category 1	0.435	0.016	1.000	0.000	1.000	1.000
Category 2	0.565	0.016	1.000	0.000	1.000	1.000

（4）加入连续结局变量的回归混合模型

加入自评抑郁得分作为结局变量，采用DCAT法，语句见框2-21。对于连续结局变量的回归混合模型，M*plus* 输出结果中会出现跨类均值相等性检验，其中会输出相关概率值。参考框2-22的输出结果。两个类别组抑郁自评得分分别为：4.54和2.90，差异显著（$p < 0.001$）。从框2-22的输出结果来看，也就是说，平均而言，生活不能自理的老人，抑郁程度要显著高于生活能够自理的老人。

框2-21　加入分类结局变量回归混合模型的M*plus* 语句

```
Title： Regression Mixture Modeling with continuous outcome variable
Data: File is older__survey. dat;
Variable: Names＝C2A C2B C2C C2D C2E C2F C2G C2H C2I C2J C2K C2L C2M C2N C2P C2Q ifold age gds;
  USEVARIABLES＝C2A-C2Q;
  MISSING are all（-9999）;
  CATEGORICAL＝C2A-C2Q;
```

```
  CLASSES=C（2）;
  AUXILIARY=gds（BCH）;
Analysis:
  TYPE=MIXTURE;
  PROCESSORS=4;
  LRTSTARTS=2 1 80 16;   ！配合tech14
PLOT: TYPE=PLOT3;
  SERIES=C2A-C2Q（*

）;
```

框2-22　加入分类结局变量回归混合模型输出结果（部分）

```
EQUALITY TESTS OF MEANS ACROSS CLASSES USING THE BCH PROCEDURE
WITH 1 DEGREE(S) OF FREEDOM FOR THE OVERALL TEST
GDS
                      Mean          S.E.
  Class 1            4.540         0.211
  Class 2            2.903         0.075
                   Chi-Square      P-Value
  Overall test      52.233         0.000
```

（5）潜类别变量作为调节变量的回归混合模型（混合回归模型）

研究不同生活自理状况类型中，不同年龄老人的抑郁水平。*C* 作为调节变量，年龄作为预测变量，自评抑郁得分作为结局变量。采用BCH法，语句见框2-23和框2-24。

框2-23　混合回归模型的*Mplus* 语句：第一步

```
Title: Step 1 in the BCH estimation for C as moderator;
DATA: File is older__survey. dat;
VARIABLE: Names＝C2A C2B C2C C2D C2E C2F C2G C2H C2I C2J C2K
C2L C2M C2N C2P C2Q ifold age gds;
  USEVARIABLES＝C2A-C2Q;
  MISSING are all（-9999）;
  CATEGORICAL＝C2A-C2Q;
  CLASSES＝C（2）;
  AUXILIARY＝gds age;
ANALYSIS: Type＝Mixture;
SAVEDATA: File＝older__bch. dat;
  Save＝bchweights;
```

框2-24　混合回归模型的*Mplus* 语句：第二步

```
Title: Step 2 in the BCH estimation for C as moderator;
Data: File is older__bch. dat;
Variable: Names＝C2A C2B C2C C2D C2E C2F C2G C2H C2I C2J C2K
C2L C2MC2N C2P C2Q gds
age W1-W2 MLC;
  USEVARIABLES＝gds age W1-W2;
  CLASSES＝C（2）;
```

```
  TRAINING＝W1-W2（BCH）;
Analysis: Type＝Mixture;
  Starts＝0;
Model:
  %OVERALL%
   gds ON age;
  %C#1%
   gds ON age;
  %C#2%
   gds ON age;
```

从框2-25的输出结果来看，生活自理不同状况的老人，年龄与抑郁水平的关系模式存在差异。

框2-25　混合回归模型输出结果（部分）

MODEL RESULTS	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Latent Class 1				
GDS ON				
AGE	1.441	0.007	218.776	0.000
AGESQ	0.152	0.076	2.003	0.045
Intercepts				
GDS	-51.384	0.457	-112.460	0.000
Residual Variances				
GDS	0.363	0.047	7.717	0.000
Latent Class 2				
GDS ON				
AGE	1.507	0.050	30.364	0.000
AGESQ	-0.224	0.753	-0.298	0.766
Intercepts				
GDS	-55.911	3.382	-16.534	0.000
Residual Variances				
GDS	0.363	0.047	7.717	0.000

（6）混合结构方程模型

构建生活自理状况与自评抑郁分之间的混合结构方程。其中因为自评抑郁属于成熟量表，为模型简化起见，直接使用总分。由此，类别潜变量实际上还是依据生活自理能力状况进行的分组。另，为了加快模型迭代收敛，代码适当降低了估计的收敛标准和精度。语句见框2-26。

框2-26　混合结构方程模型的M*plus* 语句

```
TITLE: Structural Equation Mixture Modeling
Data:  File is older__survey. dat;
Variable: Names＝C2A C2B C2C C2D C2E C2F C2G C2H C2I C2J C2K
C2L C2MC2N C2P C2Q ifold age gds;
  USEVARIABLES＝C2A-C2Q gds;
  MISSING are all（-9999）;
  CATEGORICAL＝C2A-C2Q;
  CLASSES＝C（2）;
ANALYSIS:
  TYPE＝MIXTURE;
  ALGORITHM＝INTEGRATION;
  MCONVERGENCE＝0.01;
  ITERATIONS＝30000;
  INTEGRATION IS 9;
  PROCESSORS＝4;
```

```
MODEL:
  %OVERALL%
   F BY C2A-C2Q;
   gds ON F;
  %c#1%
    [F*

1 gds];
   gds ON F;
  %c#2%
    [F*

1 gds];
   gds ON F;
OUTPUT: TECH1 TECH8;
```

从该模型输出结果来看（见框2-27），生活能够自理类型老人占85.4%，不能自理类型老人占14.6%（注意两种潜类别的次序和前面模型发生了颠倒）。F因子得分均值后者高于前者，可以印证潜类别区分的有效性。从GDS截距来看，不能自理类别组老人抑郁水平要明显高于能够自理类别组老人。不过，从抑郁对生活自理状况的斜率来看，两类老人中，二者关系模式一致，效应量也基本相当。也就是说自理状况较差，抑郁水平相应较高。不过，输出结果中，在生活自理类别组的老人，抑郁对生活自理状况斜率统计上显著，而在不能自理老人类型中则不显著。似乎在后一种类型中，可能由于适应性的缘故，生

活自理能力低对抑郁的影响有所“钝化”，造成二者关系并不增强。另外，一定程度上也可能是由于后一种类型老人样本较小造成的。

框2-27　混合结构方程模型输出结果（部分）

```
FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES
BASED ON THEIR MOST LIKELY LATENT CLASS MEMBERSHIP

Class Counts and Proportions
    Latent
    Classes
        1                 1125            0.85357
        2                  193            0.14643
CLASSIFICATION QUALITY
      Entropy                              0.863
MODEL RESULTS
                                                              Two-Tailed
                        Estimate      S.E.   Est./S.E.         P-Value
Latent Class 1
  GDS        ON
    F                    0.872       0.173     5.046            0.000
  Means
    F                   -0.049       0.033    -1.501            0.133
  Intercepts
    GDS                  2.411       0.075    31.971            0.000
Latent Class 2
    GDS        ON
      F                  0.863       0.862     1.001            0.317
    Means
      F                  0.187       0.105     1.773            0.076
    Intercepts
      GDS                7.639       0.289    26.387            0.000
```

2.7　本章小结

混合模型作为潜变量建模的发展趋势之一，到目前为止仍处在发展的初期，很多方法都还在探索中，尽管已有少数应用研究发表，但并未受到应用研究者的重视。同样地，回归混合模型作为混合模型的一个类别目前也还是个开放的研究领域，多数方法是最近三五年提出的，而且更新的方法也在不断涌现。当然，本章介绍的方法在处理实

际问题时可能存在一些问题，需要更多的方法学研究对这些方法在各种建模条件下的表现进行深入检验。

为了方便读者对本章介绍的各种方法进行比较和选择，在Asparouhov＆Muthén（2015）的基础上，表2-1汇总了本章介绍的各种方法的使用条件和简要评价。

表2-1　各种情况处理方法汇总表

适用情况		方　法	Mplus 语句：Auxiliary = ( )	评　价
结局变量	分类变量	单步法	无单独语句	直接将类别结局变量作为 LCA 的测量指标;这种做法显然会影响测量模型;纳入不同的结局变量会造成测量模型结果的差异,因此不推荐使用。
		LTB	DCAT	是处理连续结局变量最好的方法之一,推荐使用。
	连续变量	单步法	无单独语句	非正态时表现不佳。
		BCH	BCH	是处理连续结局变量最好的方法之一,在 DU3STEP 不报告结果时使用。
		稳健三步法:类别方差不等	DU3STEP	在结局变量类别内正态分布且方差不等时表现佳,但会出现类别水平顺序变化的不足。
		稳健三步法:类别方差相等	DE3STEP	在结局变量类别内正态分布,方差相等时表现佳。
		LTB	DCON	对假设前提比较敏感,当假设违反时会扭曲估计结果,不推荐使用
		PC method	E	精确性较差,不推荐实际使用
预测变量		PC method	R	结果有偏,不推荐使用。
		单步法	无单独语句	表现良好,当变量较多时使用不便。
		稳健三步法	R3STEP	表现良好,操作方便,推荐使用。
RMM		稳健三步法	无单独语句;需 2 步手动完成	表现良好,推荐使用。
		BCH	无单独语句;需 2 步手动完成	表现良好,推荐使用。

※推荐阅读※

RMM是新近发展的模型，由于建模复杂，该领域的研究相对较少，但近两年有几项研究开发了几种建模方法。其中Bakk等的系列研究及Asparouhov＆Muthén（2014）的文章是该领域的奠基之作。

Asparouhov T．＆Muthen B．（2014）．Auxiliary variables in mixture modeling：Three-step approaches using M*plus*．*Structural Equation Modeling，21*，329-341.

Bakk，Z．，Tekle，F．B．，＆Vermunt，J．K．（2013）．Estimating the association betwen latent class membership and external variables using bias adjusted three-step approaches. In T．F．Liao（Ed．），*Sociological methodology*（pp．272-311）．Thousand Oaks，CA：Sage.

Bakk，Z．and Vermunt，J．K．（2016）．Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling，23*，20-31.

Bakk，Z．，Oberski，D．L．，＆Vermunt，J．K．（2016）．Relating latent class membership to continuous distal outcomes：improving the LTB approach and a modified three-step implementation. *Structural Equation Modeling，23*，278-289.

---

(1)　这些变量在M*plus* 中统称为辅助变量auxiliary variable。

(2)　根据最大后验概率将个体分入不同的潜类别组，然后以该分组变量进行回归分析，因此得名。

(3)　这里类似参数估计的点估计，为了考虑抽样误差的影响通常采用区间估计。

(4)　因为存在分类不确定性所以抽取多个可能值作为分类误差。

(5)　自变量是分类变量（这里的潜类别变量）因变量是连续变量的回归模型等价于单因素方差分析。

(6)　ML和BCH假设连续结局变量在类别内的分布为正态分布。

3　因子混合模型

3.1　因子混合模型概述

人类的心理属性究竟是类别的还是连续的或维度的始终存在争论（e．g．，Bernstein et al，2010；Muthén，2006；Walton，Ormel＆Krueger，2011）。例如，人格的内外向性，类型论认为个体不是内向就是外向，非此即彼，不同类别间存在明确的界限。而维度观或特质论则认为内外向是个连续体，个体在这个连续体上仅存在量的差异。再如，究竟精神疾病和正常人之间存在显然的质的差异（difference in kind），还是仅在正常与变态的连续体上存在程度的不同（difference in degree）？

与此对应，在统计方法上存在两种不同的统计模型：对应维度观的是因子分析模型。因子是连续的，个体在因子上存在量的差异。类型观则对应着潜类别模型，不同的潜类别意味着存在不可忽略的类别差异。如果能有一种模型同时兼顾类型观和维度观，那么类型观和维度观之间的差异将被化解。

因子混合模型（Factor Mixture Model，FMM）就是这样一种方法，同时具有潜在类别模型和因子模型[(1)]的特点（Lubke＆Muthén，2005）。也就是说在FMM中，同时用连续的因子和间断的潜类别两种潜变量对观测数据进行建模，通过同时抽取两种潜变量来解释观测指标间的关联，即达成局部独立性。FMM既可以像LCA那样将个体划分到某个类别群体中去，同时又可以像因子分析那样考虑个体在连续因子变量上程度的差异。

3.2 因子混合模型作为一般的模型

因子混合模型的数学表达式与因子分析模型类似，只是考虑到了群体异质性，在公式上加了类别变量的下标，其意指类别特定的因子模型，公式如下：

$$
\begin{aligned}
y_{it} &= \boldsymbol{\tau}_t + \Lambda_t \boldsymbol{\eta}_{it} + \boldsymbol{\varepsilon}_{it} \\
\boldsymbol{\eta}_{it} &= \alpha_t + \zeta_{it} \\
\zeta_{it} &\sim N(0, \boldsymbol{\Psi}_t)
\end{aligned}
\tag{3.1}
$$

y 为条目，i 为个体，t 为潜类别变量， τ 为项目阈限向量（二分变量），如果是连续变量可改成截距 ν ， Λ 为因子负荷矩阵， η 为因子分向量， α 为因子均值向量； ζ 为假设正态分布的因子分残差矩阵，其均值为0，方差为 Ψ 。凡是带下标 t 的符号均表示可以估计类别特定参数，即随类别不同而变化。

在FMM示意图（图3-1）中，从类别潜变量 C 指向连续潜变量 η 的箭头表示不同潜类别间存在潜均值差异（类别特定的潜均值），对应公式中的 α_t ；由 C 指向 η 因子残差的箭头表示类别特定的方差和协方差，对应式中的 ζ_{it} ，指向因子 η 的箭头表示潜类别内部因子分存在方差变异；由 C 指向条目负荷的虚线表示因子负荷在不同类别间是变化的，对应式中的 Λ_t 。由 C 指向条目的箭头表示条目阈限（截距）是跨类别变化的，对应式中的 τ_t 。从 C 指向条目误差的箭头表示类别特定的误差，对应式中的 ε_{it} 。

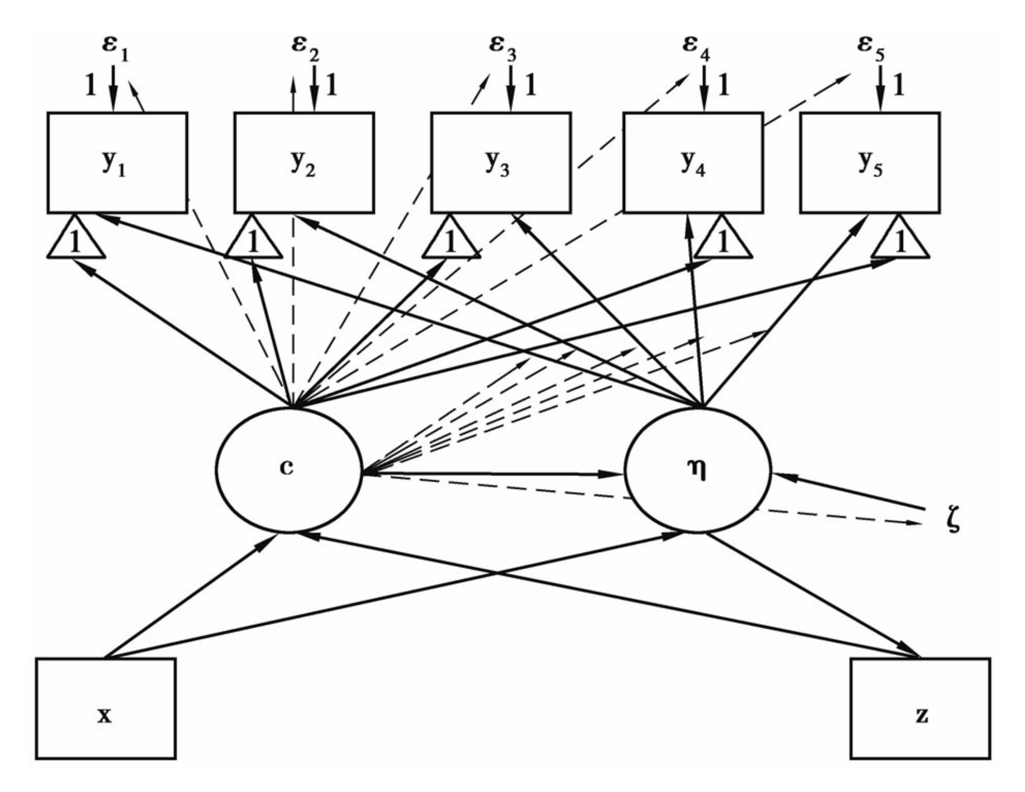


图3-1　FMM示意图（Masyn，Henderson＆Greenbaum，2010）

上述模型还可以扩展到包含前因和结果变量的模型，表达式如下：

$$
\begin{aligned}
y_{it} &= \tau_t + \Lambda_t \eta_{it} + \Gamma_t x_i + \varepsilon_{it} \\
\eta_{it} &= \alpha_t + \Gamma_t x_i + \zeta_{it}
\end{aligned}
\tag{3.2}
$$

Γ_t 为协变量x 预测 η 的类别特定的回归系数，对应图中x 指向 η 的路径。由x 指向C 的路径表示个体的类别属性受x 的影响；由C 和 η 指向Z 的路径表示个体类别属性和 η 值差异可以预测变量Z 的得分（根据Z 的分布选择不同的回归模型）。

3.3　FMM的变式

3.3.1 概述

因子模型和潜在类别模型可以作为因子混合模型的特例（3种模型的比较见表3-1）。当潜类别变量只有一个类别时就不存在群体异质性，因子混合模型简化为因子模型即上式的t＝1；当各潜类别内的方差为零时，因子混合模型简化为潜类别模型或潜在剖面模型（Lubke＆Muthén，2005）。

表3-1　3种模型的比较

模　型	特　点
潜类别模型	将外显变量间的关联解释为个体属于不同的潜类别，类别内指标间不允许相关。
因子模型	将外显变量间的关联解释为个体在连续因子上的差异，全部样本同质。
因子混合模型	将外显变量间的关联解释为个体属于不同的潜类别，同时允许指标间存在相关。

可观测的分组变量通常被看作调节变量，换句话说，就是检验某因子模型在分组变量不同水平上是否成立，如果成立说明分组变量的调节效应不显著。与之类似，类别潜变量也可视作调节变量，即检验某因子模型在不可观测的类别组中是否等值，如果模型成立说明模型具有跨潜类别的不变性，即调节效应不显著。同多组模型，存在不同水平的跨类别不变性，根据不变性的设置可以衍生出多个FMM变式。Muthén（2008）根据是否设置测量不变性以及连续潜变量的分布情况（参数vs. 非参数分布）将混合模型分成4类：混合因子分析

（Mixture Factor Analysis，MFA）、非参数混合因子分析、因子混合模型和非参数混合模型。Masyn，Henderson＆Greenbaum（2010）在一般FMM基础上，根据模型中连续潜变量和类别潜变量的特点描述了8种模型，这些模型均可看作FMM的特例或变式。下面介绍几种常见模型的具体参数设置情况和各自的特点。

3.3.2　混合因子模型（Mixture Factor Model，MFM）

在混合因子模型中，通常限定测量不变性但允许类别内方差和协方差自由估计，公式如下：

$$
\begin{aligned}
y_{it} &= \boldsymbol{\tau} + \Lambda \boldsymbol{\eta}_{it} + \boldsymbol{\varepsilon}_{it} \\
\boldsymbol{\eta}_{it} &= \alpha_t + \zeta_{it} \\
\zeta_{it} &\sim N(0, \Psi_t)
\end{aligned}
\tag{3.3}
$$

该式与FMM的一般形式相比，τ 和 Λ 参数没有下标 t，说明这些参数跨类别不变，意味着观察指标所测量的连续潜变量在各类别间有着相同的意义（见“基础篇”第7章）。进一步来说，不同潜类别个体间的差异可以通过其在 η 上的差异来比较。

在图3-2的左图中，指向因子 F 的箭头代表因子存在方差变异，对应右图中的分布和式中的 ζ_{it}。右图中的正态分布说明总体由多个正态分布的亚组混合而成，因此各潜类别内部个体间存在变异（量的差异）。潜类别的个数决定分布的个数，位置代表因子分数（α_t）的差异。

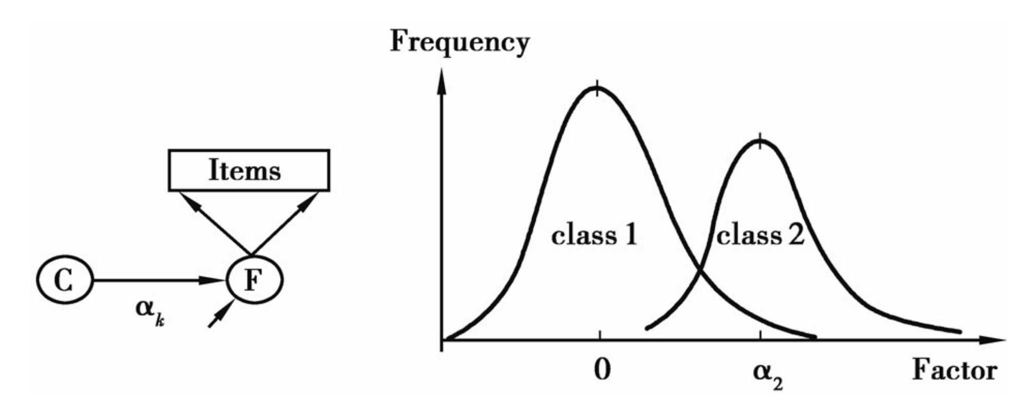


图3-2　混合因子分析模型示意图

混合因子模型与传统的因子模型相比，释放了多元正态分布的限定，整体分布是由多个正态分布混合而成的非特定的连续型分布。混合因子模型可以理解为将非正态的整体分布分解为多个亚组，这些亚组符合或接近正态分布，并且因子模型在每个亚组内满足严格测量不变性，因此不同潜类别个体的差异可以通过其在 η 上的差异来比较。如果将传统的因子模型称作参数模型，那么混合因子模型也称作半参数因子分析（Semi-parametric Factor Analysis，SP-FA；Masyn，Henderson＆Greenbaum，2010）。

3.3.3　潜类别因子分析（latent Class Factor Analytic，LCFA）

当测量模型满足严格不变性时，即限定项目阈限（thresholds）和因子负荷跨类别等值，而只允许因子均值跨类别变化（同MFM的区别在于不允许类别内方差和协方差自由估计），称作潜类别因子分析模型，其表达式如下：

$$
\begin{aligned}
y_{it} &= \boldsymbol{\tau} + \Lambda \boldsymbol{\eta}_{it} + \varepsilon_{it} \\
\boldsymbol{\eta}_{it} &= \alpha_t
\end{aligned}
\tag{3.4}
$$

该式与MFM相比，缺失的 ζ 和 Ψ 说明因子方差和协方差固定为0，即类别内不存在变异，模型示意图如下：

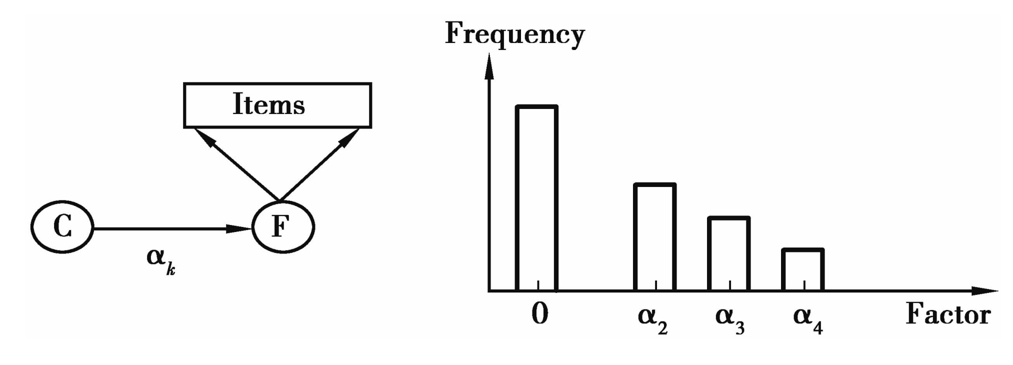


图3-3 LCFA模型示意图

左图中，C 指向F 的路径表示因子均值跨类别变化，图中缺少从C 指向条目的路径，说明条目截距或阈限跨类别等值，或者说个体的类别属性仅由个体在因子上的得分差异决定；缺少从C 指向因子负荷路径的虚线说明因子负荷跨类别不变。指向因子F的箭头消失说明因子不存在方差变异。右图中的矩形表示各潜类别因子分，说明类别内不存在变异即同类别内个体有相同的因子分（但不意味着有相同的观测分），换句话说，连续潜变量不存在参数分布，因此该模型也称作非参数MFM（non-parametric MFM；Muthén，2008）和非参数因子分析模型（Non-parametric Factor Analysis，NP-FA；Masyn，Henderson＆Greenbaum，2010）。矩形的个数取决于潜类别的个数，矩形的高度代表类别内个体的数量或比例，矩形所在位置表示因子得分。

前两个模型满足严格测量不变性，条目所测量的连续潜变量 η 在不同类别具有相同的意义，所以不管个体属于哪个潜类别均可在潜变量 η 上进行差异比较（Masyn，Henderson＆Greenbaum，2010）。

3.3.4 半参数因子混合模型（Semi-parametric Factor Mixture Model，SP-FMM）

半参数因子混合模型（SP-FMM）是FMM和半参数因子模型（SP-FM）结合后的模型，具有两者各自的特点。通俗地说，SP-FMM就是使用一个潜类别变量将总体分成几个异质的亚组或类别（对应图3-4的右图中左右两簇分布），然后再在每个类别内单独做一个SP-FM（对应3-4左图中各簇分布内相互重叠的分布图）。可见该模型存在两个潜在类别变量，第一个潜类别变量C^g 用于描述总体异质性，第二个潜类别C^{sp} 用于解释类别内的非正态性。在SP-FMM中，类别间测量模型不再限定严格不变性（在图3-4的右图中用间断的水平线表示），所以不同类别的因子模型测量不同的潜变量 η_t ，也意味着不同类别个体不能在同一个 η 上进行差异比较。然而SP-FMM具有SP-FM的特征，即每个类别内部满足SP-FM。由于SP-FM要求满足严格测量不变性，所以类别内部的亚组间可以在各自的η t上进行差异比较。

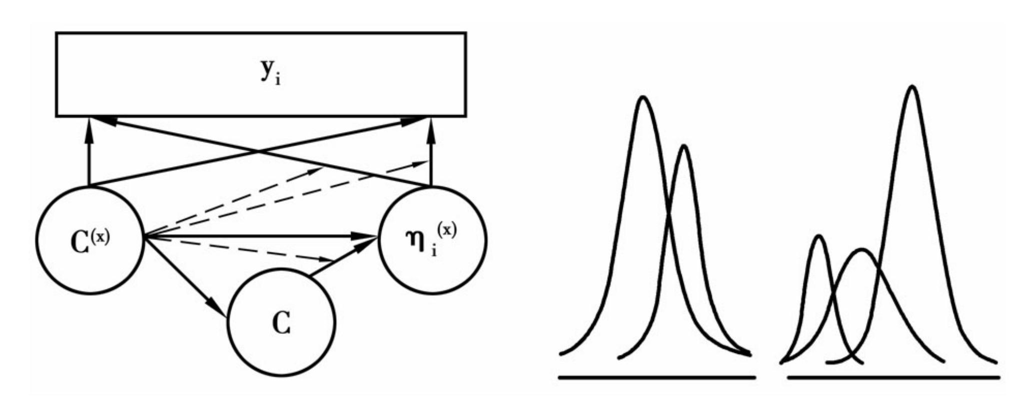


图3-4 半参数因子混合模型示意图

3.3.5 非参数因子混合模型（Non-parametric Factor Mixture Model，NP-FMM）

非参数因子混合模型是FMM和非参数因子模型（NP-FM）结合后的模型，同时具有两者的特点。NP-FMM与SP-FMM非常类似，差异仅在于各异质群体内部的各亚组的因子是否存在方差变异，如果存在方差变异则用分布表示，否则用矩形，其他特点同SP-FMM。

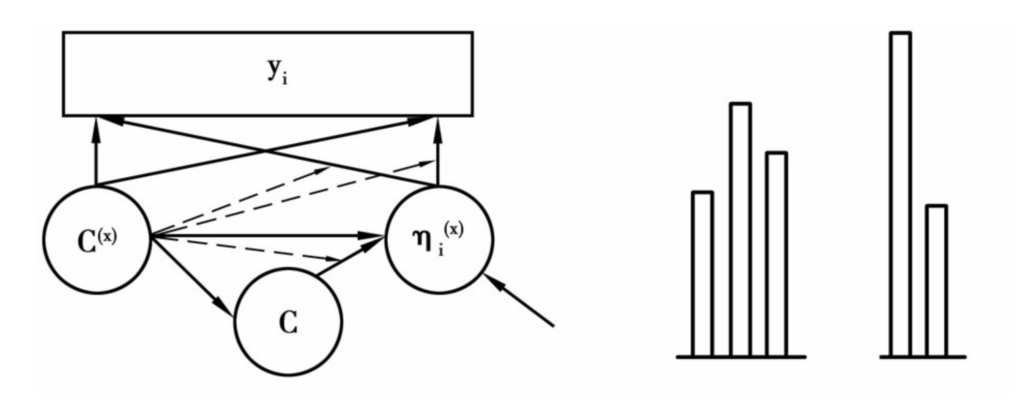


图3-5 非参数因子混合模型示意图

3.3.6 其他变式

如果设定因子负荷和因子协方差跨类别不变，项目阈限（或截距）跨类别变化，公式如下（FMM-1）：

$$
\begin{aligned}
y_{it} &= \tau_t + \Lambda\eta_{it} + \varepsilon_{it} \\
\eta_{it} &= \zeta_i \\
\zeta_{it} &\sim N(0,\Psi)
\end{aligned}
\tag{3.5}
$$

此时的模型因子均值为0，所以在示意图上（图3-6）由C指向F的箭头消失了，同时多了由C指向指标的箭头，说明项目阈限跨类别变化即类别属性基于项目反应获得。

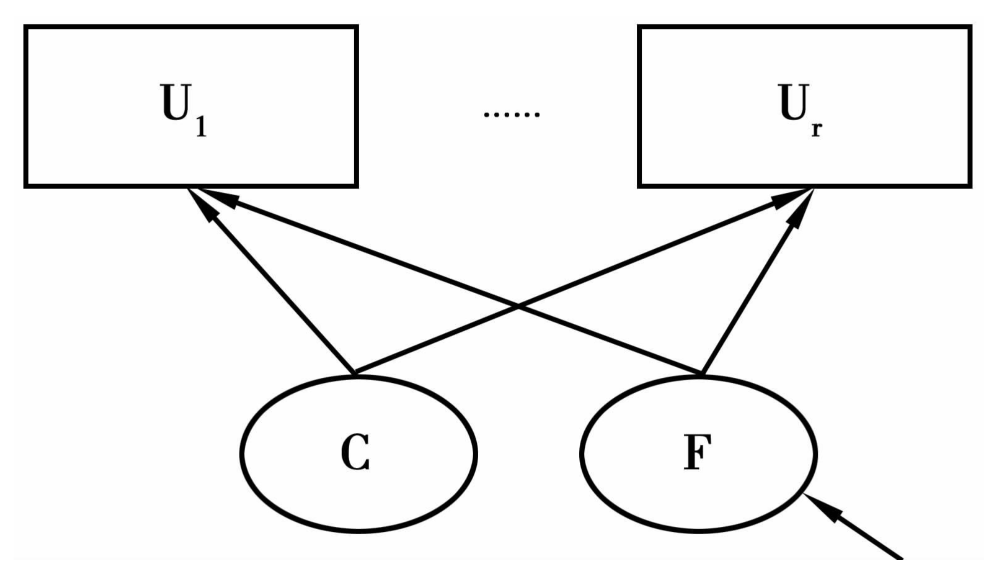


图3-6　FMM-3示意图

在上式的基础上允许协方差矩阵跨类别变化，表明在类别内部个体存在量的差异，表达式如下（FMM-2）：

$$
\begin{aligned}
y_{it} &= \boldsymbol{\tau}_t + \Lambda \boldsymbol{\eta}_{it} + \varepsilon_{it} \\
\boldsymbol{\eta}_{it} &= \zeta_{it} \\
\zeta_{it} &\sim N(0, \Psi_t)
\end{aligned}
\tag{3.6}
$$

进一步设定因子负荷跨类别变化，即得到类似于一般FMM的模型，表达式如下（FMM-3）：

$$
\begin{aligned}
y_{it} &= \boldsymbol{\tau}_t + \Lambda_t \boldsymbol{\eta}_{it} + \varepsilon_{it} \\
\boldsymbol{\eta}_{it} &= \zeta_{it} \\
\zeta_{it} &\sim N(0, \Psi_t)
\end{aligned}
\tag{3.7}
$$

表3-2总结了上述6种FMM相关模型的公式及参数设定要求，所有模型间的区别在于测量不变性和模型中因子参数分布是否不同。有些模型在实际应用中拟合研究数据存在困难（Clark，2010，下面的示例结果也存在拟合不良的问题），但分析思路同MCFA，即从限定最松散的模型开始依次增加限定参数，通过比较统计指数差异确定拟合最优模型。

表3-2　几种常见FMM变式参数设置小结

模　型	表达式	不变性参数	变化参数	其他设定
FMM	$y_{it}=\boldsymbol{\tau}_t+\Lambda_t\eta_{it}+\varepsilon_{it}$ $\eta_{it}=\alpha_t+\zeta_{it}$ $\zeta_{it}\sim N(0,\Psi_t)$	无	α_t Λ_t $\boldsymbol{\tau}_t$ ζ_{it} Ψ_t	无
LCFA 或 NP-FM	$y_{it}=\boldsymbol{\tau}+\Lambda\eta_{it}+\varepsilon_{it}$ $\eta_{it}=\alpha_t$	$\boldsymbol{\tau}$ Λ	α_t	$\zeta=0$ $\Psi=0$
MFM 或 SP-FM	$y_{it}=\boldsymbol{\tau}+\Lambda\eta_{it}+\varepsilon_{it}$ $\eta_{it}=\alpha_t+\zeta_{it}$ $\zeta_{it}\sim N(0,\Psi_t)$	$\boldsymbol{\tau}$ Λ	α_t ζ_{it} Ψ_t	无
FMM-1	$y_{it}=\boldsymbol{\tau}_t+\Lambda\eta_{it}+\varepsilon_{it}$ $\eta_{it}=\zeta_i$ $\zeta_{it}\sim N(0,\Psi)$	Λ ζ_i Ψ	$\boldsymbol{\tau}_t$	$\alpha=0$
FMM-2	$y_{it}=\boldsymbol{\tau}_t+\Lambda\eta_{it}+\varepsilon_{it}$ $\eta_{it}=\zeta_{it}$ $\zeta_{it}\sim N(0,\Psi_t)$	Λ	$\boldsymbol{\tau}_t$ ζ_{it} Ψ_t	$\alpha=0$
FMM-3	$y_{it}=\boldsymbol{\tau}_t+\Lambda_t\eta_{it}+\varepsilon_{it}$ $\eta_{it}=\zeta_{it}$ $\zeta_{it}\sim N(0,\Psi_t)$	无	$\boldsymbol{\tau}_t$ Λ_t ζ_{it} Ψ_t	$\alpha=0$

3.4　FMM的分析过程

FMM作为新发展的方法在实际研究中的应用并不多，到目前为止，尚未形成统一的分析步骤。但就目前为数不多的应用研究来看（e．g．，Bernstein et al，2010；Clark，2010；Muthén，2006；Walton Ormel＆Krueger，2011），多数研究采用FA、LCA和FMM3种模型同时比较的思路，即同时使用3种模型拟合数据，然后通过比较统计指标选择最优模型（混合模型评价和比较的指数见第1章的相关内容）。下面按照多数研究者的做法，总结了FMM的分析步骤：

①拟合因子模型，逐步增加因子个数以确定最优因子结构；

②拟合潜在类别模型，逐步增加类别个数以确定最佳潜类别模型；

③根据因子模型和潜类别模型确定FMM的类别数和因子模型；

④逐渐减少类别个数和因子个数以获取最佳FMM模型；

⑤比较FMM与因子模型和潜类别模型，以获得最佳的拟合模型。

关于上述步骤有3点需要说明。第一，前两步的顺序并不重要，可以先做FA后做LCA，也可先做LCA后做FA。第二，LCA和FA最佳模型的类别或因子个数应该是随后FMM类别和因子个数的上限。因为如果存在程度上的差异，LCA会使用更多的类别来解释，所以当这种差异在FMM中使用连续的因子来解释时，相应的类别数应减少，至少不应增加。第三，潜类别个数和因子个数的减少应遵循先类别后因子的顺序。

3.5　实例分析

3.5.1　PTSD结构的FMM分析

下面使用PTSD的数据演示FMM的分析过程。如前所述，由于FMM作为因子模型和LCA的一般形式，通常在执行FMM分析时会执行因子模型和LCA模型，用以比较3类竞争模型（e．g．，Bernstein et al，2010；Walton，Ormel＆Krueger，2011），下面采用上面总结的步骤演示FMM的过程。具体来说，先比较FA、LCA和LCFA对数据的拟合情况，然后再比较不同FMM变式之间对数据拟合的差异。

FA、LCA和LCFA拟合结果呈现在表3-3中，FA和LCA的M*plus* 语句可参见“基础篇”第5章和本书第1章的相关内容，LCFA的M*plus* 语句呈现在框3-1中。

3.5.2　结果解释

（1）确立最佳模型

下面主要演示分析过程，因此不考虑模型选择的理论意义。因子模型比较了4个竞争模型，结果表明King的四因子模型的信息指数最小，通过比较我们选择King四因子模型作为因子模型中的最佳模型。

LCA使用PTSD五点计分数据，所以表中结果与表1-7的结果有所不同。信息指数随抽取类别数目的增加而减少并未出现转折，所以在此根据信息指数很难得到合适的模型；另外两个统计指标LMR和aLRT提示保留3个类别较为合适，综合考虑后选择三类别LCA为最佳模型。

因子模型和LCA模型的最佳模型分别确定，随后分别执行2～4个类别以King四因子结构为测量模型的FMM即LCFA。如前所述，FMM存在多个变式，不同的变式对FMM参数的设置不同，所以同为FMM分析，不同的参数限定会得到不同的结果。此处，我们先执行截距和因子负荷跨类别等值的FMM，即LCFA或NP-FM，随后再比较不同变式的FMM差异。

从表3-3的模型拟合指数来看，四类别FMM各拟合指数均比三类别要低一些，同时LMR和BLRT统计结果并没有明确的提示，然而4个类别中有1个类别比例仅为1.79%，所以3个类别的结果较为合理。

（2）3种模型之间的比较

3种分析模型的最佳模型分别确立：King四因子测量模型，三类别LCA模型和三类别四因子的FMM。下面比较3种模型并从中选择最佳模型。从拟合指数的角度来看，King的四因子模型各信息指数最小，其次是三类别的LCA模型，而三类别FMM模型的指数最大。因此从信息指数角度来说，用因子模型拟合PTSD数据最合适，说明PTSD潜在结构是连续的而非间断的，即个体在PTSD上只存在量的差异而非质的差别。这里需要说明的是，FMM模型存在不同变式，此处设定的FMM模型为相对严格的模型，下面将逐步释放这些限制以比较其他模型的拟合情况。

（3）M*plus* 设置[(2)]

FMM在M*plus*的设定非常类似于多组CFA的设置，使用“％overall％”表示整体情况，凡是其后的语句均表示所有类别适用。使用“％c＃1％”表示潜类别1，其后的语句仅适用于潜类别1。有时需要比较不同类别因子均值差异，可以通过在OUT*P* UT后加TECH4获得模型中所有连续潜变量的均值。

M*plus* 默认最后一个类别为参照类别组，所以该组的因子均值默认为0，其他类别的因子均值自由估计。M*plus* 默认因子负荷、截距或阈限和观测指标的测量误差跨类别不变，即公式中的τ、Λ 和指标测量误差保持等值，其他参数跨类别自由估计。LCFA的M*plus* 语句呈现在框3-1中。

（4）FMM的变式结果

FMM各变式的拟合结果汇总在表3-4中。在LCFA的基础上允许类别内存在方差变异的MFM或SP-FM（M*plus* 语句见框3-3）拟合结果优于LCFA模型。BIC减少了519，说明类别内存在方差变异。然而MFM抽取3个类别时LMR和BLRT检验不显著，提示抽取2个类别优于3个类别，同时Entropy＝0.694提示较高的分类错误，因此检验了抽取2个类别的MFM。两个类别的MFM信息指数有细微升高，而LMR和BLRT检验显著，分类精确性指标Entropy也有所提高，因此保留2类别MFM为最佳模型。

表3-3　因子模型和潜类别模型拟合指数汇总表

Model	k	ln(L)	AIC	BIC	aBIC	Entropy	LMR	BLRT	Class Proportions
CFA									
DSM	54	−13820.017	27748.035	27981.743	27810.321	–			
King	**57**	**−13756.893**	**27627.787**	**27874.479**	**27693.533**	**–**			
Simms	57	−13760.972	27635.943	27882.636	27701.690	–			
Elhai	61	−13764.259	27638.519	27876.555	27701.959	–			
LCA									
One-class	34	−15107.454	30282.909	30430.059	30322.126	–	–	–	
Two-class	52	−14102.360	28308.719	28533.772	28368.698	.890	0	0	.5756.4245
Three-class	**70**	**−13841.322**	**27822.644**	**28125.599**	**27903.385**	**.897**	**.007**	**.008**	**.4823.3907.1270**
Four-class	88	−13741.099	27658.198	28039.056	27759.701	.901	.706	.706	.4875.3571.1304.025
LCFA：King's model									
Two-class	52	−14102.360	28308.719	28533.772	28368.698	.890	.001	.002	.4286.5714
Three-class	**57**	**−13859.042**	**27832.084**	**28078.776**	**27897.830**	**.880**	**.012**	**.000**	**.3878.1389.4733**
Four-class	62	−13780.096	27684.191	27952.523	27755.705	.881	.003	.000	.0179.1857.3429.4536

表3-4　因子模型和潜类别模型拟合指数汇总表

Model	k	ln(L)	AIC	BIC	aBIC	Entropy	LMR	BLRT	Class Proportions
FMM：									
NP-FM/LCFA-3c	57	−13859.042	27832.084	28078.776	27897.830	.880	.012	.000	.3878.1389.4733
SP-FM/MFA-3c	83	−13517.178	27200.357	27559.576	27296.093	.694	.463	.111	.1696.4554.3750
SP-FM/MFA-2c	68	−13557.420	27250.841	27545.141	27329.275	.779	.000	.000	.5875.4125
FMM-1-3C	93	−13524.679	27235.358	27637.856	27342.628	.926	.1318	.000	.6875.0929.2196
FMM-1-2C	75	−13596.277	27342.554	27667.149	27429.063	.961	.000	.000	.7854.2146
FMM-2-2C	85	−13525.586	27221.172	27589.047	27319.216	.830	.029	.000	.3750.6250
FMM-3-2c	123	−12890.000	26026.000	26558.336	26167.874	.945	.2209	.000	.2940.7060
FMM-2c	127	−12890.000	26034.000	26583.648	26180.488	.945	.2398	000	.3036.6964

框3-1　三类别LCFA的M*plus* 语句表

```
TITLE: LCFA or NP-FMof PTSD with 3 class and King model
DATA: FILE IS PTSD. dat;
VARIABLE: NAMES ARE x y0 y1-y17;
    USEVARIABLES ARE y1-y17;
    CLASSES＝c（3）;
ANALYSIS: TYPE＝MIXTURE;
    ESTIMATOR＝MLR;
    STARTS＝20010;
    Processor＝2;
MODEL: %overall%！这里用%overall%表示整体情况，下面的语句表示所有类别都适用。
    f1 by y1-y5; ！因子模型，因子负荷等值；
    f2 by y6-y7;
```

```
    f3 by y8-y12;
    f4 by y13-y17;
    f1-f4@0; ! 根据LCFA设定因子方差为0;
    f1 with f2-f4@0; ! 根据LCFA设定因子协方差为0;
    f2 with f3-f4@0;
    f3 with f4@0;
  %c#1%
     [f1-f4*

];! 类别1的因子均值自由估计;
  %c#2%
     [f1-f4*

];! 类别2的因子均值自由估计;
    ! %c#3%软件默认最后一个类别为参照组,因子均值固定为0;
OUTPUT: TECH4 TECH11 tech14;
SAVEDATA: file=PTSD-LCFA-king-3c. txt;
    save=cprob;
```

框3-2　三类别LCFA的模型结果

下面分别给出 3 个类别特定的测量参数。

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Latent Class 1 ！下面是 class1 的非标准化因子负荷，由于设定了因子负荷和截距等值，这些参数在 3 个类别里是相等的。				
F1 BY				
Y1	1.000	0.000	999.000	999.000
Y2	1.010	0.144	6.999	0.000
Y3	1.438	0.141	10.189	0.000
Y4	1.231	0.135	9.090	0.000
Y5	1.441	0.160	9.021	0.000
F2 BY				
Y6	1.000	0.000	999.000	999.000
Y7	1.216	0.088	13.871	0.000
F3 BY				
Y8	1.000	0.000	999.000	999.000
Y9	1.003	0.115	8.732	0.000
Y10	1.231	0.134	9.202	0.000
Y11	1.132	0.148	7.659	0.000
Y12	0.801	0.130	6.154	0.000
F4 BY				
Y13	1.000	0.000	999.000	999.000
Y14	1.026	0.111	9.240	0.000
Y15	1.063	0.111	9.537	0.000
Y16	1.142	0.135	8.462	0.000
Y17	1.066	0.120	8.878	0.000
F1 WITH ！根据 LCFA，因子协方差设定为 0。				
F2	0.000	0.000	999.000	999.000
F3	0.000	0.000	999.000	999.000
F4	0.000	0.000	999.000	999.000
F2 WITH				
F3	0.000	0.000	999.000	999.000
F4	0.000	0.000	999.000	999.000
F3 WITH				
F4	0.000	0.000	999.000	999.000
Means				
F1	0.823	0.096	8.598	0.000
F2	1.025	0.100	10.262	0.000
F3	0.661	0.081	8.182	0.000
F4	0.889	0.086	10.327	0.000
Intercepts				
Y1	1.831	0.061	30.074	0.000
Y2	1.456	0.057	25.558	0.000
Y3	1.527	0.058	26.233	0.000
Y4	2.063	0.070	29.569	0.000
Y5	1.505	0.060	25.037	0.000
Y6	1.506	0.046	33.053	0.000

Y7	1.393	0.040	35.185	0.000
Y8	1.606	0.055	29.155	0.000
Y9	1.447	0.053	27.077	0.000
Y10	1.247	0.045	27.827	0.000
Y11	1.324	0.060	21.926	0.000
Y12	1.369	0.064	21.257	0.000
Y13	1.460	0.064	22.648	0.000
Y14	1.510	0.057	26.278	0.000
Y15	1.726	0.057	30.542	0.000

Y16	1.437	0.056	25.451	0.000
Y17	2.147	0.078	27.406	0.000
Variances ! 根据 **LCFA**,因子方差设定为 0。				
F1	0.000	0.000	999.000	999.000
F2	0.000	0.000	999.000	999.000
F3	0.000	0.000	999.000	999.000
F4	0.000	0.000	999.000	999.000
Residual Variances				
Y1	1.041	0.068	15.263	0.000
Y2	1.107	0.075	14.789	0.000
Y3	0.943	0.067	14.002	0.000
Y4	1.189	0.073	16.376	0.000
Y5	1.088	0.083	13.102	0.000
Y6	0.856	0.059	14.560	0.000
Y7	0.761	0.054	14.077	0.000
Y8	0.893	0.068	13.095	0.000
Y9	0.992	0.077	12.807	0.000
Y10	0.674	0.104	6.508	0.000
Y11	0.789	0.065	12.083	0.000
Y12	1.073	0.089	12.009	0.000
Y13	1.050	0.083	12.608	0.000
Y14	0.986	0.070	14.027	0.000
Y15	0.989	0.068	14.589	0.000
Y16	1.036	0.068	15.195	0.000
Y17	1.267	0.075	16.860	0.000
Latent Class 2				
F1 BY				
Y1	1.000	0.000	999.000	999.000
Y2	1.010	0.144	6.999	0.000
Y3	1.438	0.141	10.189	0.000
Y4	1.231	0.135	9.090	0.000
Y5	1.441	0.160	9.021	0.000
F2 BY				
Y6	1.000	0.000	999.000	999.000
Y7	1.216	0.088	13.871	0.000
F3 BY				
Y8	1.000	0.000	999.000	999.000
Y9	1.003	0.115	8.732	0.000
Y10	1.231	0.134	9.202	0.000
Y11	1.132	0.148	7.659	0.000
Y12	0.801	0.130	6.154	0.000
F4 BY				

Y13	1.000	0.000	999.000	999.000
Y14	1.026	0.111	9.240	0.000
Y15	1.063	0.111	9.537	0.000
Y16	1.142	0.135	8.462	0.000
Y17	1.066	0.120	8.878	0.000
F1 WITH				
F2	0.000	0.000	999.000	999.000
F3	0.000	0.000	999.000	999.000
F4	0.000	0.000	999.000	999.000
F2 WITH				
F3	0.000	0.000	999.000	999.000
F4	0.000	0.000	999.000	999.000
F3 WITH				
F4	0.000	0.000	999.000	999.000
Means				
F1	1.378	0.181	7.632	0.000
F2	1.427	0.158	9.045	0.000
F3	1.921	0.204	9.418	0.000
F4	1.688	0.197	8.573	0.000
Intercepts				
Y1	1.831	0.061	30.074	0.000
Y2	1.456	0.057	25.558	0.000
Y3	1.527	0.058	26.233	0.000
Y4	2.063	0.070	29.569	0.000
Y5	1.505	0.060	25.037	0.000
Y6	1.506	0.046	33.053	0.000
Y7	1.393	0.040	35.185	0.000
Y8	1.606	0.055	29.155	0.000
Y9	1.447	0.053	27.077	0.000
Y10	1.247	0.045	27.827	0.000
Y11	1.324	0.060	21.926	0.000
Y12	1.369	0.064	21.257	0.000
Y13	1.460	0.064	22.648	0.000
Y14	1.510	0.057	26.278	0.000
Y15	1.726	0.057	30.542	0.000
Y16	1.437	0.056	25.451	0.000
Y17	2.147	0.078	27.406	0.000
Variances				
F1	0.000	0.000	999.000	999.000
F2	0.000	0.000	999.000	999.000
F3	0.000	0.000	999.000	999.000
F4	0.000	0.000	999.000	999.000

Residual Variances				
Y1	1.041	0.068	15.263	0.000
Y2	1.107	0.075	14.789	0.000
Y3	0.943	0.067	14.002	0.000
Y4	1.189	0.073	16.376	0.000
Y5	1.088	0.083	13.102	0.000
Y6	0.856	0.059	14.560	0.000
Y7	0.761	0.054	14.077	0.000
Y8	0.893	0.068	13.095	0.000
Y9	0.992	0.077	12.807	0.000
Y10	0.674	0.104	6.508	0.000
Y11	0.789	0.065	12.083	0.000
Y12	1.073	0.089	12.009	0.000
Y13	1.050	0.083	12.608	0.000
Y14	0.986	0.070	14.027	0.000
Y15	0.989	0.068	14.589	0.000
Y16	1.036	0.068	15.195	0.000
Y17	1.267	0.075	16.860	0.000
Latent Class 3				
F1 BY				
Y1	1.000	0.000	999.000	999.000
Y2	1.010	0.144	6.999	0.000
Y3	1.438	0.141	10.189	0.000
Y4	1.231	0.135	9.090	0.000
Y5	1.441	0.160	9.021	0.000
F2 BY				
Y6	1.000	0.000	999.000	999.000
Y7	1.216	0.088	13.871	0.000
F3 BY				
Y8	1.000	0.000	999.000	999.000
Y9	1.003	0.115	8.732	0.000
Y10	1.231	0.134	9.202	0.000
Y11	1.132	0.148	7.659	0.000
Y12	0.801	0.130	6.154	0.000
F4 BY				
Y13	1.000	0.000	999.000	999.000
Y14	1.026	0.111	9.240	0.000
Y15	1.063	0.111	9.537	0.000
Y16	1.142	0.135	8.462	0.000
Y17	1.066	0.120	8.878	0.000
F1 WITH				
F2	0.000	0.000	999.000	999.000

```
    F3                 0.000      0.000    999.000    999.000
    F4                 0.000      0.000    999.000    999.000
F2       WITH
    F3                 0.000      0.000    999.000    999.000
    F4                 0.000      0.000    999.000    999.000
F3       WITH
    F4                 0.000      0.000    999.000    999.000
```

Means ! **Mplus 默认设定最后一个类别为参照类别，所以前两个类别的因子均值计算，最后一个类别的因子均值固定为** 0。

```
    F1                 0.000      0.000    999.000    999.000
    F2                 0.000      0.000    999.000    999.000
    F3                 0.000      0.000    999.000    999.000
    F4                 0.000      0.000    999.000    999.000
Intercepts
    Y1                 1.831      0.061     30.074      0.000
    Y2                 1.456      0.057     25.558      0.000
    Y3                 1.527      0.058     26.233      0.000
    Y4                 2.063      0.070     29.569      0.000
    Y5                 1.505      0.060     25.037      0.000
    Y6                 1.506      0.046     33.053      0.000
    Y7                 1.393      0.040     35.185      0.000
    Y8                 1.606      0.055     29.155      0.000
    Y9                 1.447      0.053     27.077      0.000
    Y10                1.247      0.045     27.827      0.000
    Y11                1.324      0.060     21.926      0.000
    Y12                1.369      0.064     21.257      0.000
    Y13                1.460      0.064     22.648      0.000
    Y14                1.510      0.057     26.278      0.000
    Y15                1.726      0.057     30.542      0.000
    Y16                1.437      0.056     25.451      0.000
    Y17                2.147      0.078     27.406      0.000
Variances
    F1                 0.000      0.000    999.000    999.000
    F2                 0.000      0.000    999.000    999.000
    F3                 0.000      0.000    999.000    999.000
    F4                 0.000      0.000    999.000    999.000
Residual Variances
    Y1                 1.041      0.068     15.263      0.000
    Y2                 1.107      0.075     14.789      0.000
    Y3                 0.943      0.067     14.002      0.000
    Y4                 1.189      0.073     16.376      0.000
    Y5                 1.088      0.083     13.102      0.000
```

Y6	0.856	0.059	14.560	0.000
Y7	0.761	0.054	14.077	0.000
Y8	0.893	0.068	13.095	0.000
Y9	0.992	0.077	12.807	0.000
Y10	0.674	0.104	6.508	0.000
Y11	0.789	0.065	12.083	0.000
Y12	1.073	0.089	12.009	0.000
Y13	1.050	0.083	12.608	0.000
Y14	0.986	0.070	14.027	0.000
Y15	0.989	0.068	14.589	0.000
Y16	1.036	0.068	15.195	0.000
Y17	1.267	0.075	16.860	0.000

框3-3　三类别SP-FM或MFM模型的M*plus* 语句

```
TITLE: SP-FMor MFMof PTSD with 3 class and king model
DATA: FILE IS PTSD. dat;
VARIABLE: NAMES ARE x y0 y1-y17;
    USEVARIABLES ARE y1-y17;
    CLASSES＝c（3）;
ANALYSIS: TYPE＝MIXTURE;
    ESTIMATOR＝MLR;
    STARTS＝200 10;
    Processor＝2;
MODEL:   %overall%
    f1 by y1-y5;
    f2 by y6-y7;
    f3 by y8-y12;
    f4 by y13-y17;
    f1-f4; ! 因子方差估计，但限定跨类别不变。
    f1 with f2-f4; ! 因子方差估计，但限定跨类别不变。
    f2 with f3-f4;
```

```
    f3 with f4;
  %c#1%
    [f1-f4*
];! 均值自由估计
  %c#2%
    [f1-f4*
];
OUTPUT: TECH4 TECH11 tech14;
SAVEDATA: file=PTSD-MFM-king-3c.txt;
    save=cprob;
```

设置负荷，方差和协方差不变，截距或阈限跨类别变化的FMM-1抽取3个类别时的LMR参数为0.1318，提示减少一个类别的模型拟合更好。随后估计抽取2个类别的FMM-1，结果发现该模型各信息指数均较3类别模型有所升高，ΔBIC≈30＞10，而LMR和BLRT检验显著，分类精确性指标Entropy也有所提高，因此保留2类别MFM为最佳模型。FMM-1与LCFA相比，释放了截距不变和因子方差、协方差等于零的限制（尽管因子方差协方差被限定跨类别不变，但与设定为0相比仍做了退让）所以从整个模型的限制上来看，FMM-1相对较为自由，模型自由参数较多（93 vs. 57），各信息指数也相对较低，ΔBIC≈440＞10，提示FMM-1拟合数据更理想。FMM-1的M*plus* 语句见框3-4，限于篇幅，结果未呈现。

框3-4　FMM-1模型的M*plus* 语句

```
TITLE: FMM-1 of PTSD with 3 class and king model
DATA: FILE IS PTSD. dat;
VARIABLE: NAMES ARE x y0 y1-y17;
    USEVARIABLES ARE y1-y17;
    CLASSES＝c（3）；
ANALYSIS: TYPE＝MIXTURE;
    ESTIMATOR＝MLR;
    STARTS＝200 10;
    processor＝2;
MODEL: %overall%
    f1 by y1-y5; ! 负荷等值;
    f2 by y6-y7;
    f3 by y8-y12;
    f4 by y13-y17;
    f1-f4; ! 方差设定不变;
    f1 with f2-f4; ! 协方差设定不变;
    f2 with f3-f4;
    f3 with f4;
     [f1-f4@0]；! 为了模型识别，因子均值固定为0；或者设定条目截距为0。
  %c#1%
     [y1-y17]；! 截距设定自由估计;
  %c#2%
     [y1-y17]；
  %c#3%
     [y1-y17]；
```

```
OUTPUT：TECH4 TECH11 tech14；
SAVEDATA：file＝FMM-1-king-3c. txt；save＝cprob；
```

在FMM-1抽取2个类别模型的基础上允许因子方差和协方差跨类别变化的FMM-2模型，各信息指数较FMM-1有所下降，其中ΔBIC≈78＞10，提示限定因子方差和协方差跨类别不变是不合适的。FMM-2的M*plus* 语句见框3-5，由于篇幅限制，结果并未呈现。

框3-5　FMM-2模型的M*plus* 语句

```
TITLE：FMM-1 of PTSD with 2 class and king model
DATA：FILE IS PTSD. dat；
VARIABLE：NAMES ARE x y0 y1-y17；
   USEVARIABLES ARE y1-y17；
   CLASSES＝c（3）；
ANALYSIS：TYPE＝MIXTURE；
    ESTIMATOR＝MLR；
    STARTS＝200 10；
    processor＝2；
MODEL：％overall％
    f1 by y1-y5；
    f2 by y6-y7；
    f3 by y8-y12；
    f4 by y13-y17；
     [f1-f4@0]；
```

```
  %c#1%
    f1-f4; ! 方差跨类别自由估计;
    f1 with f2-f4; ! 协方差跨类别自由估计;
    f2 with f3-f4;
    f3 with f4;
    [y1-y17];
  %c#2%
    f1-f4;
    f1 with f2-f4;
    f2 with f3-f4;
    f3 with f4;
    [y1-y17];
  %c#3%
    f1-f4;
    f1 with f2-f4;
    f2 with f3-f4;
    f3 with f4;
    [y1-y17];
OUTPUT: TECH4 TECH11 tech14;
SAVEDATA: file=FMM-2-king-3c.txt;  save=cprob;
```

FMM-3的M*plus*语句见框3-6，各拟合指数结果进一步下降ΔBIC≈1031，分类精确性进一步提高至0.945，LMR和BLRT的结果冲突，因此2个类别的FMM-3模型拟合较FMM-2要好，提示限定负荷跨类别不变并不合适。完全自由的FMM模型结果表明，拟合指数有所提高，

ΔBIC≈25＞10，其他结果与FMM-3一致，仅从统计指标角度考虑，FMM-3模型最为理想。FMM的M*plus* 语句呈现在框3-7中。

框3-6　FMM-3模型的M*plus* 语句

```
TITLE: FMM-1 of PTSD with 2 class and king model
DATA: FILE IS PTSD. dat;
VARIABLE: NAMES ARE x y0 y1-y17;
    USEVARIABLES ARE y1-y17;
    CLASSES＝c（2）；！3个类别存在估计问题，选择2个类别
ANALYSIS: TYPE＝MIXTURE;
    ESTIMATOR＝MLR;
    STARTS＝20010;
    processor＝2;
MODEL: %overall%
    f1 by y2-y5；！因子负荷跨类别自由估计；
    f2 by y7;
    f3 by y9-y12;
    f4 by y14-y17;
     [f1-f4@0]；
  %c#1%
    f1 by y2-y5；！因子负荷跨类别自由估计；
    f2 by y7;
    f3 by y9-y12;
    f4 by y14-y17;
    f1-f4;
```

```
    f1 with f2-f4;
    f2 with f3-f4;
    f3 with f4;
    [y1-y17];
  %c#2%
    f1 by y2-y5;
    f2 by y7;
    f3 by y9-y12;
    f4 by y14-y17;
    f1-f4;
    f1 with f2-f4;
    f2 with f3-f4;
    f3 with f4;
    [y1-y17];
  %c#3%
    f1 by y2-y5;
    f2 by y7;
    f3 by y9-y12;
    f4 by y14-y17;
    f1-f4;
    f1 with f2-f4;
    f2 with f3-f4;
    f3 with f4;
    [y1-y17];
OUTPUT: TECH4 TECH11 tech14;
SAVEDATA: file=FMM-3-king-3c. txt; save=cprob;
```

框3-7　FMM模型的M*plus* 语句

```
TITLE: FMMof PTSD with 2 class and king model
DATA: FILE IS PTSD. dat;
VARIABLE: NAMES ARE x y0 y1-y17;
    USEVARIABLES ARE y1-y17;
    CLASSES＝c（2）;
ANALYSIS: TYPE＝MIXTURE;
    ESTIMATOR＝MLR;
    STARTS＝200 10;
    processor＝2;
MODEL: %overall%
    f1 by y1-y5;
    f2 by y6-y7;
    f3 by y8-y12;
    f4 by y13-y17;
  %c#1%
     [f1-f4*

];
    f1 by y1-y5;
    f2 by y6-y7;
    f3 by y8-y12;
    f4 by y13-y17;
    f1-f4;
    f1 with f2-f4;
```

```
    f2 with f3-f4;
    f3 with f4;
     [y1-y17];
  %c#2%
    f1 by y1-y5;
    f2 by y6-y7;
    f3 by y8-y12;
    f4 by y13-y17;
    f1-f4;
    f1 with f2-f4;
    f2 with f3-f4;
    f3 with f4;
     [y1-y17];
OUTPUT: TECH4 TECH11 tech14;
SAVEDATA: file=FMM-3-king-3c.txt;
    save=cprob;
```

3.6 本章小结

在前一章潜类别分析的基础上，本章的内容进一步拓宽了潜类别分析模型的范围。FMM将连续和类别潜变量统一到一个模型中，是对因子模型和潜类别模型的拓展，使群体异质性和个体程度差异得以同时考虑。相对于因子模型和LCA来说，FMM还是个新鲜事物，分析过程尚未完善，但无疑FMM将成为新宠。

※推荐阅读※

Muthén（2008）的文章根据是否满足测量不变性及连续潜变量的分布情况（参数vs. 非参数分布）将混合模型分成四类：混合因子分析、非参数混合因子分析、因子混合模型和非参数混合模型。Masyn，Henderson＆Greenbaum（2010）的文章在一般FMM基础上，根据模型中连续潜变量和类别潜变量的特点描述了8种模型，这些模型均可看作FMM的特例或变式。Walton，Ormel＆Krueger（2011）的文章通过比较FA、LCA和FMM分析了青少年外向行为的结构。

Masyn，K. ，Henderson，C. ，＆Greenbaum，P. （2010）. Exploring the latent structures of psychological constructs in social development using the Dimensional-Categorical Spectrum. *Social Development* ，19，470-493.

Muthén，B. （2008）. Latent variable hybrids. In G. R. Hancock，＆K. M. Samuelson（Eds. ），*Advances in latent variable mixture models* （pp. 1-24）. Charlotte，NC：Information Age Publishing.

Walton，K. E. ，Ormel，J. ，＆Krueger，R. F. （2011）. The Dimensional Nature of Externalizing Behaviors in Adolescence：Evidence from a Direct Comparison of Categorical，Dimensional，and Hybrid Models. *Journal ofAbnormal Child Psychology* ，39，553-561.

(1)　由于FMM带有FA的特征，所以存在探索和验证两种不同的FMM，本章主要论述验证性的FMM，探索性的FMM看参考McLachlan和Peel（2000）专著。

(2)　本章示例采用M*plus* 6.11，如果采用M*plus* 7.4结果会有所差异。

4　潜增长曲线模型

4.1 潜增长模型概述

在心理学、教育学和社会学等社会科学领域，以及在医学和自然科学领域，对事物发展进程的研究也常是研究者的兴趣点。例如，在教育学领域，教育学者对学生阅读能力随年级增加而增长的情况的研究；在心理学领域，发展心理学家对人格特质随年龄增长而变化的趋势的研究；在医学领域，研究者观察癌细胞增殖变化的时间进程。诸如此类，都涉及对事物随时间变化过程的刻画。

对于此类问题，不同学科发展出不同的方法。在心理学领域，研究者在结构方程模型框架内发展出潜增长曲线模型（Latent Growth Curve Models，LGCM；Kaplan，2000）或潜曲线模型（Latent Curve Model，LCM；e．g．，Meredith＆Tisak，1990；Bollen＆Curran，2006）。统计和生物统计学领域发展了随机系数模型（e．g．，Laird＆Ware，1982）。教育学领域则提出了多水平模型（Multilevel Modeling，MLM；e．g．，Bryk＆Raudenbush，1987）或分层线性模型（Hierarchical Linear Modeling，HLM）。尽管这些方法在形式上有所差异，但在统计原理上则大同小异。LCM和MLM相比各有优点和不足，在有些条件下两者等同（e．g．，Curran，2003），但LGCM在功能上要灵活一些，而且在多数结构方程建模软件上均可实现（e．g．，AMOS，LISREL，M*plus*，EQS）[(1)]。

即使在SEM内部，LGCM也有很多种不同的称谓。例如，潜在轨迹模型（Latent Trajectory Model，LTM；e．g．，Curran＆Hussong，2003），潜增长曲线模型（LGCM；Kaplan，2000），潜成长模型（Latent Growth Models，LGM；McArdle＆Bell，2000）或潜变量成长曲线模型（Latent Variable Growth Curve Modeling；e．g．，

Duncan，Duncan，＆Strycker，2006）。不同的研究者偏爱不同的称谓，甚至同一研究者在不同的场合也使用不同的名称（e．g．，McArdle＆Bell，2000）。本书统一采用潜增长曲线模型（e．g．，Kaplan，2000）[(2)]。

本章主要介绍SEM框架下的LGCM，在此之前首先解释多水平框架下如何描述个体的成长轨迹。由于多水平模型是对一般线性回归的扩展，所以通过介绍多水平模型可以帮助我们更好地理解个体成长轨迹的特点，进而更容易理解SEM下的LGCM。

4.2　潜增长模型的原理

4.2.1　多水平模型对个体成长轨迹的描述

纵向数据或重复测量数据（3次测量以上）可以勾画出单个个体的成长轨迹，如图4-1所示。

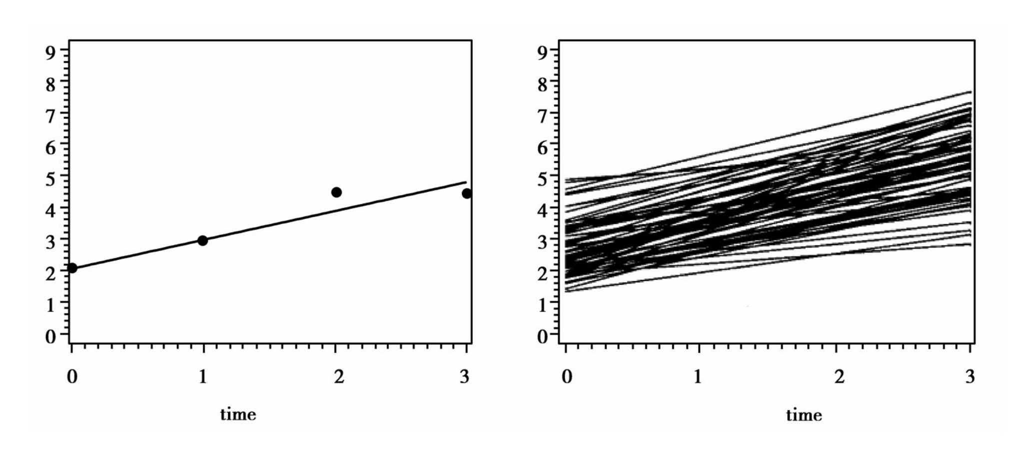


图4-1　个体成长轨迹

图4-1以测量时间点为横坐标，变量得分为纵坐标，将个体在4次重复测量上的得分呈现在坐标图上。左图为单个个体4次测量的得分（实心黑点）和趋势线（直线），右图为多个个体的成长轨迹（即多个个体的趋势线）。根据回归的知识，我们可以为每个个体建立各自的回归方程：

$$y_{it} = \alpha_{it} + \beta_{it} \lambda_{it} + \varepsilon_{it} \quad (4.1)$$

其中i ＝1，2，…，N，表示个体，N为样本量；t ＝1，2，…，T，表示测量时间点，T为总次数。y_{it}表示个体i在时间点t时的得分，α_i为个体i成长轨迹的截距；β_i为个体i成长轨迹的斜率，λ_t表示时间点t时的得分；ε_{it}为个体i在时间点t时的残差项。将i和t替换成具体的数字可产生一系列的方程。

然而，个体间和个体内部（不同时间段间的差异）均存在差异，表现在初始值和成长轨迹或斜率上。如果个体间不存在初始值的差异，那么所有的个体只有一个共同的起点；如果个体间不存在斜率的差异，那么所有个体只有一条成长轨迹线。多数情况下，个体在初始值和成长轨迹或斜率上是存在差异的，可以通过如下两个方程描述这种差异：

$$\alpha_i = \mu_\alpha + \zeta_{\alpha i}$$
$$\beta_i = \mu_\beta + \zeta_{\beta i}$$

α_i和β_i含义同上，μ_α和μ_β分别表示全部个体截距和斜率的均值即总均值，由于每个个体具有相同的μ_α和μ_β，因此也称固

定系数；$\zeta_{\alpha i}$ 和 $\zeta_{\beta i}$ 分别表示个体截距和斜率与对应的总均值间的差异，每个个体均有一个特定的值，因此也称作随机系数。

将上式合并后组成联合方程：

$$y_{it} = (\mu_\alpha + \lambda_t \mu_\beta) + (\zeta_{\alpha i} + \lambda_t \zeta_{\beta i} + \varepsilon_{it}) \quad (4.2)$$

上述模型需要满足如下假设条件：所有误差或残差服从均值为0，方差为 σ^2 的正态分布，且误差间、误差与斜率和截距间不相关。

如果将重复测量看作水平1的数据，个体作为水平2数据，即不同时间点的测量嵌套于个体，上述模型则变成多水平模型（此处为两水平模型）。在多水平模型框架内解释个体成长更加直观同时也易于理解，下面介绍SEM如何处理个体成长数据。

4.2.2　LGCM描述个体成长

在LGCM中，通过假设的潜在发展轨迹（Growth Trajectory）来刻画重复测量间的变化类型。下面以不包含解释变量的LGCM模型，即无条件LGCM来阐述LGCM刻画发展轨迹的特点。需要说明的是，下面仍然从非技术的角度进行介绍，SEM框架内的代数解释可参见Duncan，Duncan＆Strycker（2006）或Bollen＆Curran（2006）。

在图4-2中，通过4次重复测量的观测指标定义两个潜变量，分别代表潜截距因子 α 和潜斜率因子 β 。α 和 β 各存在均值和方差两个参数，LGCM就是通过这两个参数描述组间和组内差异的。具体来说，截距因子的均值（相当于 μ_α ）描述平均的初始状态，截距因子

的方差（相当于$\zeta_{\alpha i}$）表示个体在特定时间点间的差异程度或离散程度，其值越大说明个体间初始差异越明显。斜率因子的均值表示时间点间的平均增长率（相当于μ_{β}），而斜率因子的方差（相当于$\zeta_{\beta i}$）反映个体间增长率差异的大小，方差越大表明个体间发展轨迹差异越明显。

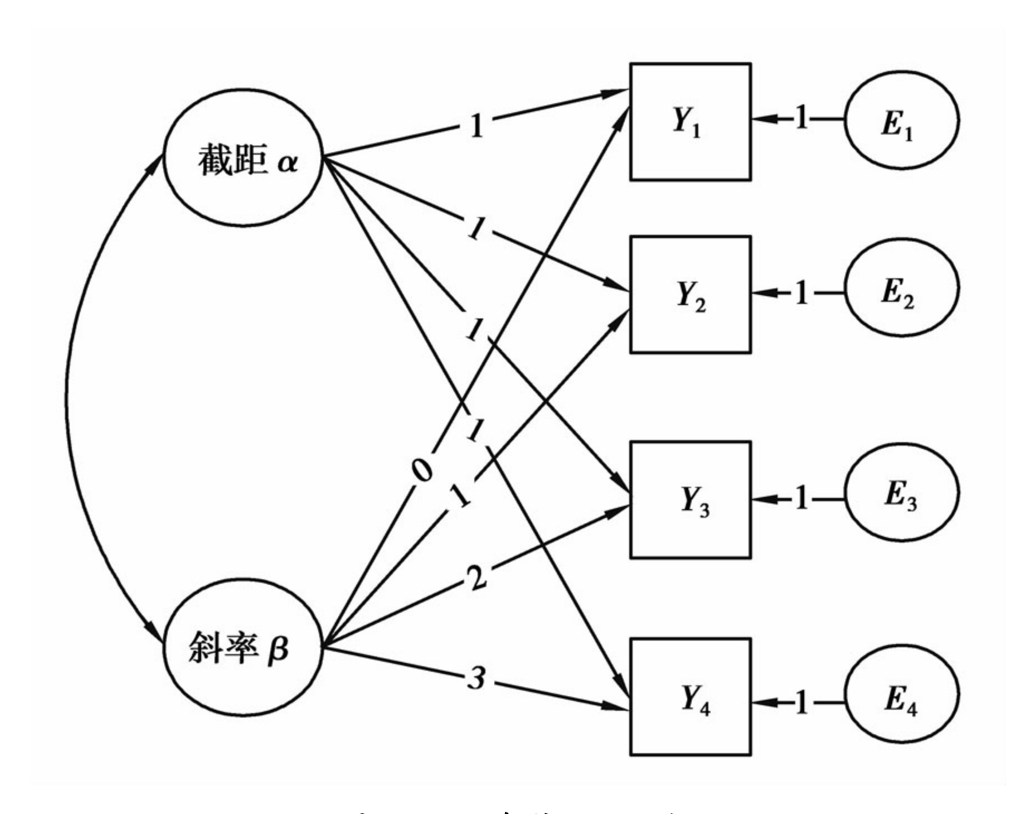


图4-2　无条件LGCM示例

重复测量在截距因子上的负荷固定为1，表示每次测量时的截距不变（测量等值），斜率因子上的负荷表示时间效应（Meredith＆Tisak，1990），可以设定具体的值，也可以允许自由估计，不同的设定表示不同的时间函数或成长类型。例如，上述4次测量分别为第一次，6个月，12个月和24个月时的测量，可以将4次时间设定为等距的间隔$t_1=0$，$t_2=6$，$t_3=12$和$t_4=24$，或$t_1=0$，$t_2=1$，

t_3 =2和t_4 =4，表示时间函数为线性即线性成长模型。α 和β 之间的双箭头表示两个因子间的相关，用于说明个体变化的截距和变化斜率间的关系。正相关表明较高的初始值对应正向增长斜率；负相关表明，较高的初始值对应负向增长斜率。

4.2.3　潜增长模型与多水平模型比较

不少研究者从技术和非技术的角度论证了多水平模型和LGCM在处理线性纵向数据时的等价性（e．g．，Curran，2003；Raudenbush，2001）。多水平模型设计之初就是用于处理嵌套结构的数据。比如，学生嵌套于班级，班级嵌套于学校，学校嵌套于学区。嵌套数据本身具有其特殊性，由于低水平的数据嵌套于高水平数据，同水平的数据之间存在某种关联而非完全独立。例如，来自相同班级的学生具有更多的共同性，不同学校的学生身上带有其学校的特性。重复测量数据也存在嵌套的特点，即单个时间点的测量嵌套于个体，不同时间点的测量存在相关，用多水平模型处理重复测量数据是其分内之事。多水平模型相比LGCM更容易设置，且较少存在参数估计问题（Bauer，2003；Curran，2003）。多水平模型作为传统回归模型的扩展仍然保留了传统回归的不足（Bauer，2003）。首先，MLM没有考虑测量误差。其次，不能直接检验中介效应。第三，不能对整个模型进行拟合检验。

这些不足恰好是SEM的特点和强项。除了上述特点外，LGCM还可以将模型灵活地扩展，①包含前因和结果（outcome）变量；②包含类别潜变量分析群体异质性（即增长混合模型，Growth Mixture

Modeling，GMM；王孟成，毕向阳，叶浩生，2014）；③进行多水平增长模型分析。

当然，LGCM也存在一些不足之处，主要表现在嵌套数据或非独立数据违反了SEM参数估计的前提假设，即局部独立性假设。SEM主要通过将水平1的时间作为固定值纳入因子负荷矩阵来估计参数，而在多水平模型中时间是作为预测变量来处理的（Curran，2003）。

4.3　线性增长模型拓展

4.3.1　非线性增长模型

上述模型描述的增长过程或方式是线性的，即相等单位的时间变化对应相等单位的结果变量的变化。但有时事物的发展并非完全按照线性的进程展开，可能出现减速或加速的非线性增长模式，此时用线性增长模型则不能确切地刻画事物的发展轨迹，应该考虑非线性增长模型。下面描述的是最常见的非线性增长曲线模型，其他类型的非线性增长模型，如logistic回归，Gompertz和Richards函数型增长模型等更多的信息及M*plus* 实现请参见Grimm＆Ram（2009）。

（1）二次函数型（Quadratic Function）

最常用的非线性增长模型如图4-3所示，在一般线性增长模型的基础上添加了二次函数潜因子，用以刻画非线性发展进程。如同线性增长一样，通过为二次函数潜因子设置具体二次函数的负荷来实现对曲线发展轨迹的刻画。具体来说，就是通过为斜率因子设定不同的时间参数来实现，即将线性增长因子的负荷设定成等距的时间参数，将二次函数增长因子的负荷设定成二次函数型时间参数（线性时间参数的平方）。在图4-3中，线性增长因子的负荷设定为$t_1=0$，$t_2=1$，$t_3=2$和$t_4=4$；二次函数增长因子的负荷设定为$t_1=0$，$t_2=1$，$t_3=4$和$t_4=9$。依次类推，可以按照相似的方式设置三次函数增长模型或更高次的增长模型（Bollen＆Curran，2006）。

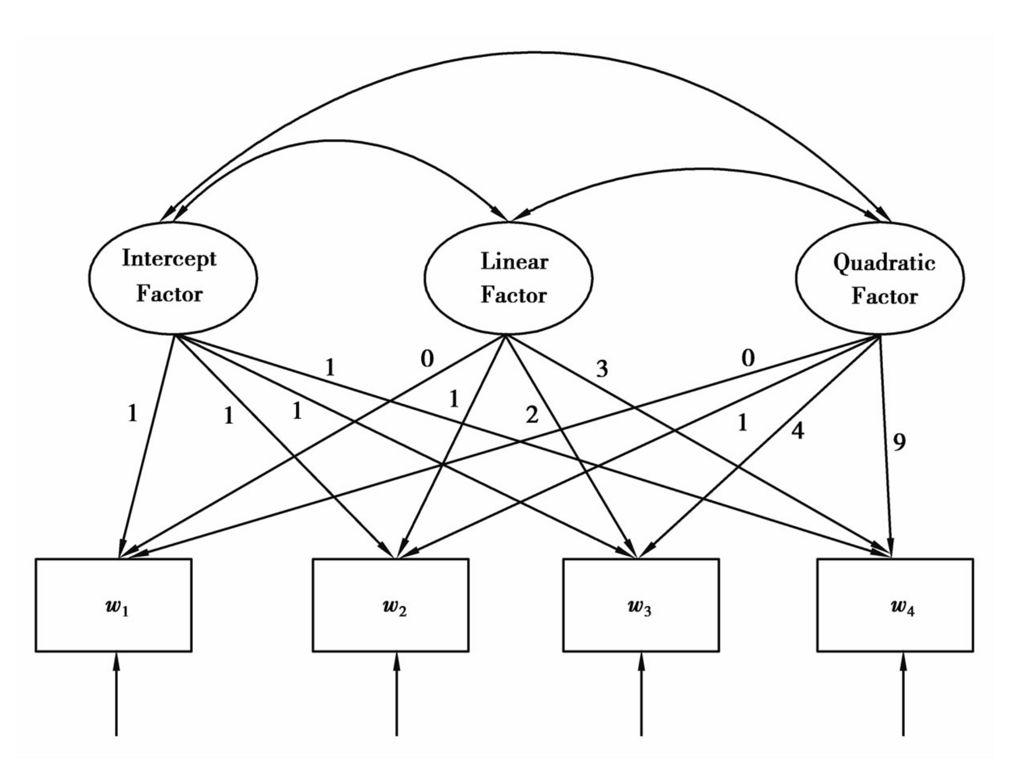


图4-3　二次函数增长模型示意图

同线性增长模型一样，下面给出了二次函数增长模型的方程，据此类推，可以设置三次函数（Cubic Function）或更高次函数的增长

方程。

$$y_{it} = \alpha_i + \beta_{1i}\lambda_t + \beta_{2i}\lambda_t^2 + \varepsilon_{it} \tag{4.3}$$

λ_t^2 为时间值的平方，β_{2i} 为曲线增长的斜率，其他符号同线性增长模型。

个体间和个体内部在初始值和成长轨迹或斜率上的差异通过如下方程来表达：

$$\alpha_i = \mu_\alpha + \zeta_{\alpha i}$$

$$\beta_{1i} = \mu_{\beta_1} + \zeta_{\beta_1 i}$$

$$\beta_{2i} = \mu_{\beta_2} + \zeta_{\beta_2 i}$$

将上述公式合并：

$$y_{it} = (\mu_\alpha + \mu_\beta \lambda_t + \mu_{\beta_2}\lambda_t^2) + (\zeta_{\alpha i} + \lambda_t \zeta_{\beta_1 i} + \lambda_t^2 \zeta_{\beta_2 i} + \varepsilon_{it}) \tag{4.4}$$

第一个括号内为固定效应系数，所有个体相同，第二个括号内为随机效应系数，反映个体间的差异。

（2）时间参数自由估计模型

有时候根据测量时间为模型设定的时间参数并不能很好地刻画增长模式，Meredith和Tisak（1990）提出在估计增长因子时将第一和第二次的负荷分别设定为0和1，随后的时间负荷允许模型从数据直接估计以确定增长类型。这种做法带着探索的味道，让数据本身决定增长模式。

图4-4是一个自由估计时间参数LGCM模型的路径图，该图与一般的LGCM路径图没有太大区别，仅有的差异在于第2次测量之后的时间参数未事先设定（* 号表示），而允许自由估计。

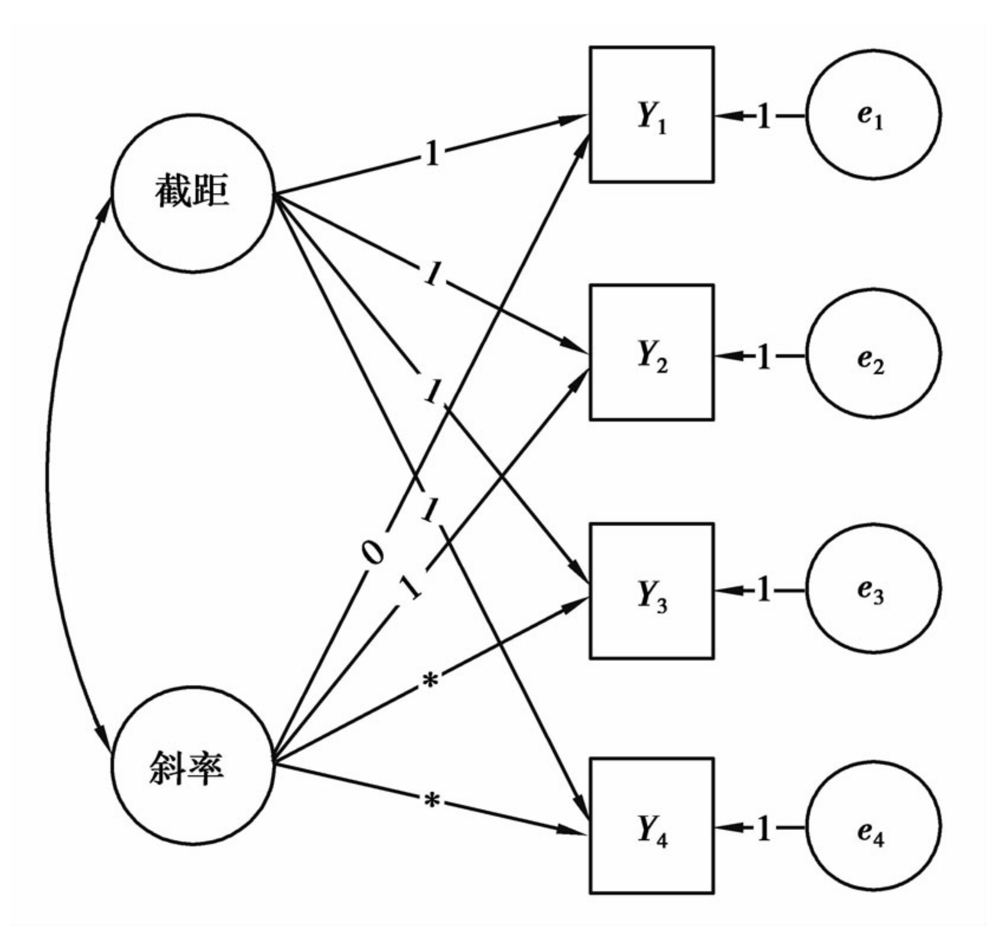


图4-4　自由估计时间参数的LGCM

4.3.2 包含协变量的LGCM

无条件的LGCM只单纯地描述了某一现象随时间而变化的进程。很多时候，事物不是孤立存在的，总是与其他事物存在千丝万缕的联系，其他事物会对研究对象的发展过程产生影响，所以要精确地刻画研究对象的发展轨迹还需要考虑或控制其他潜在因素对研究变量发展轨迹的影响。

（1）协变量类型的确定

线性增长模型包含协变量的模型称作有条件的LGCM。通常，LGCM包含的协变量可以分为两类：时间变化协变量（time-varying covariates）和时间不变协变量（time-invariant covariates）。常见的时间不变的协变量有性别、种族等，这些信息在整个重复测量过程中不会变化，测量一次即可获得；时间变化的协变量如年龄，本身随时间变化而变化，需要多次测量。

将协变量作为时间变化的协变量还是不变的协变量，应该从理论假说和实际操作两个方面来考虑。比如，某项研究准备考察外向性人格特质对大学新生入学适应过程的影响。研究者从新生入学的时候开始，每个月测量一次，连续测量6次。在这个例子中，将外向性特质作为时间变化的协变量处理，还是作为时间不变的协变量处理呢？在理论上人格特质假定是相对稳定的，在相对较短的时间内是不变的（如数年）。在本例中，间隔的时间为1个月，而且整个研究周期也相对较短（半年左右），所以将其作为时间不变的协变量处理是合适的。但

如果测量的时间间隔很长而且追踪的周期也很长，比如每间隔3年时间进行一次测量，连续追踪30年，则应该考虑将人格特质做时间变化的协变量处理，因为在时间太长的跨度内即使假设相对稳定的人格特质也是会发生变化的。

（2）时间不变协变量的LGCM

时间不变的协变量常常对发展进程产生影响。例如，前面所举的大学生入学适应的例子，不同生源地（城市vs．农村）的学生的入学适应过程可能存在很大的差异。有时研究者想考察不同人口学变量对潜在增长过程的影响，或者研究者想控制某些时间不变的协变量对潜在增长过程的影响，这些情况都需要将协变量纳入模型考虑，图4-5是一个包含时间不变协变量的LGCM路径图。

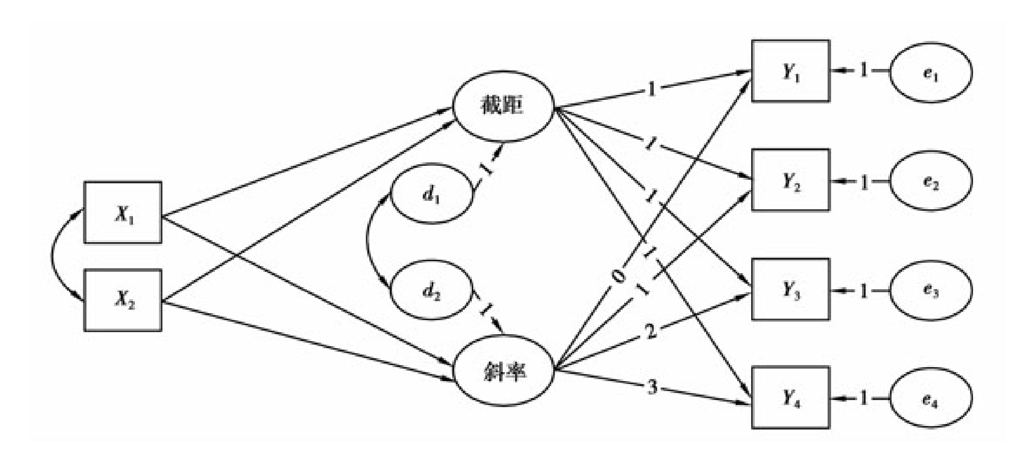


图4-5　带有时间不变协变量的LGCM

（3）时间变化协变量的LGCM

图4-6是一个包含时间变化协变量的LGCM模型路径图。由于协变量随时间变化，所以在研究变量测量时，协变量也同时测量，每个时间点的协变量只影响同时测量的研究变量，不同时间点的协变量之间，以及与潜因子之间都存在相关。

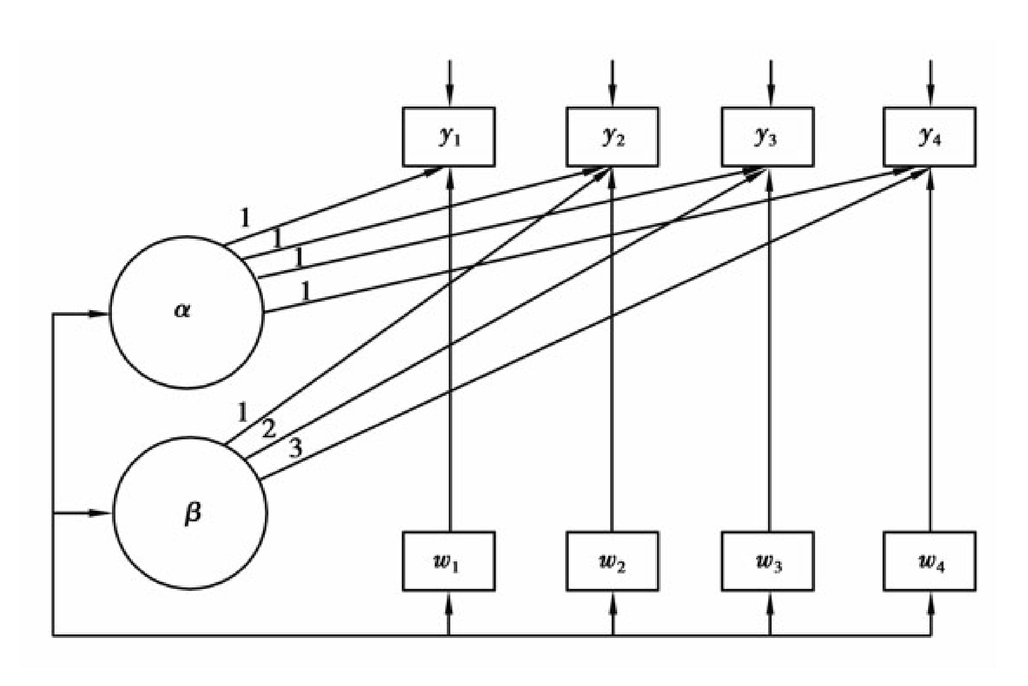


图4-6　包含时间变化协变量的LGCM路径图

协变量指向结果变量的回归系数，反映了控制结果变量潜在增长过程后的效应。

（4）同时包含时间变化和不变协变量的LGCM

在实际研究中，模型常常同时包含时间变化和时间不变两种协变量。如图4-7就是一个同时包含时间变化和时间不变协变量的LGCM，该图与4-6的区别在于多了两个时间不变的协变量。

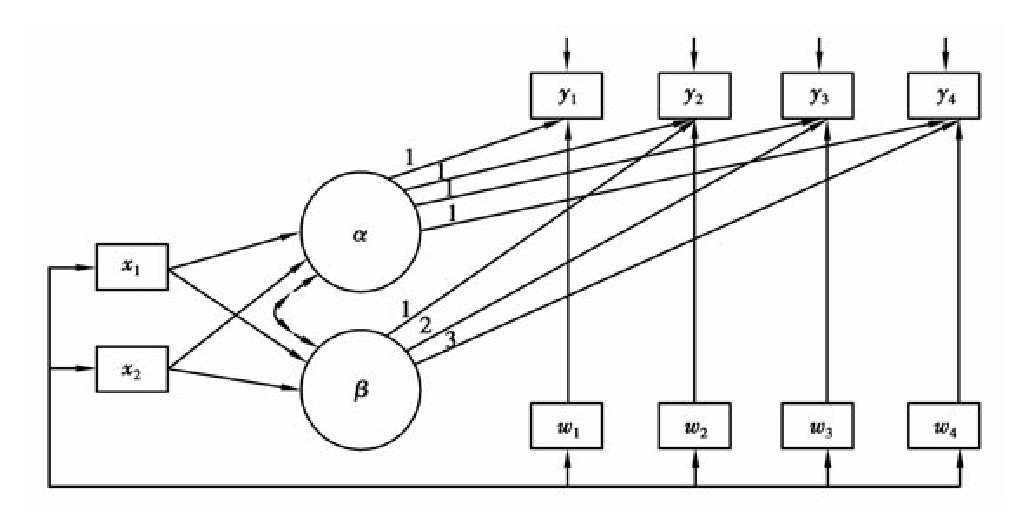


图4-7　同时包含时间变化和时间不变协变量的LGCM路径图

4.3.3　多变量增长曲线模型

上述带有时间变化协变量的LGCM将随时间变化的协变量纳入模型，控制了这些变量对研究变量潜在增长轨迹的影响，这种做法仅将随时间变化的协变量作为观测显变量处理，并没有考虑协变量自身随时间变化的特点。因此有必要在考察研究变量增长轨迹的同时，使用同样的思路处理协变量的潜在增长轨迹，在潜变量水平上控制协变量对研究变量发展进程的影响。将协变量当作单独的LGCM过程，同时考察两个LGCM之间的关系，此时的增长曲线模型称作多变量增长曲线模型（Multivariate Latent Growth CurveModel），如图4-8所示。

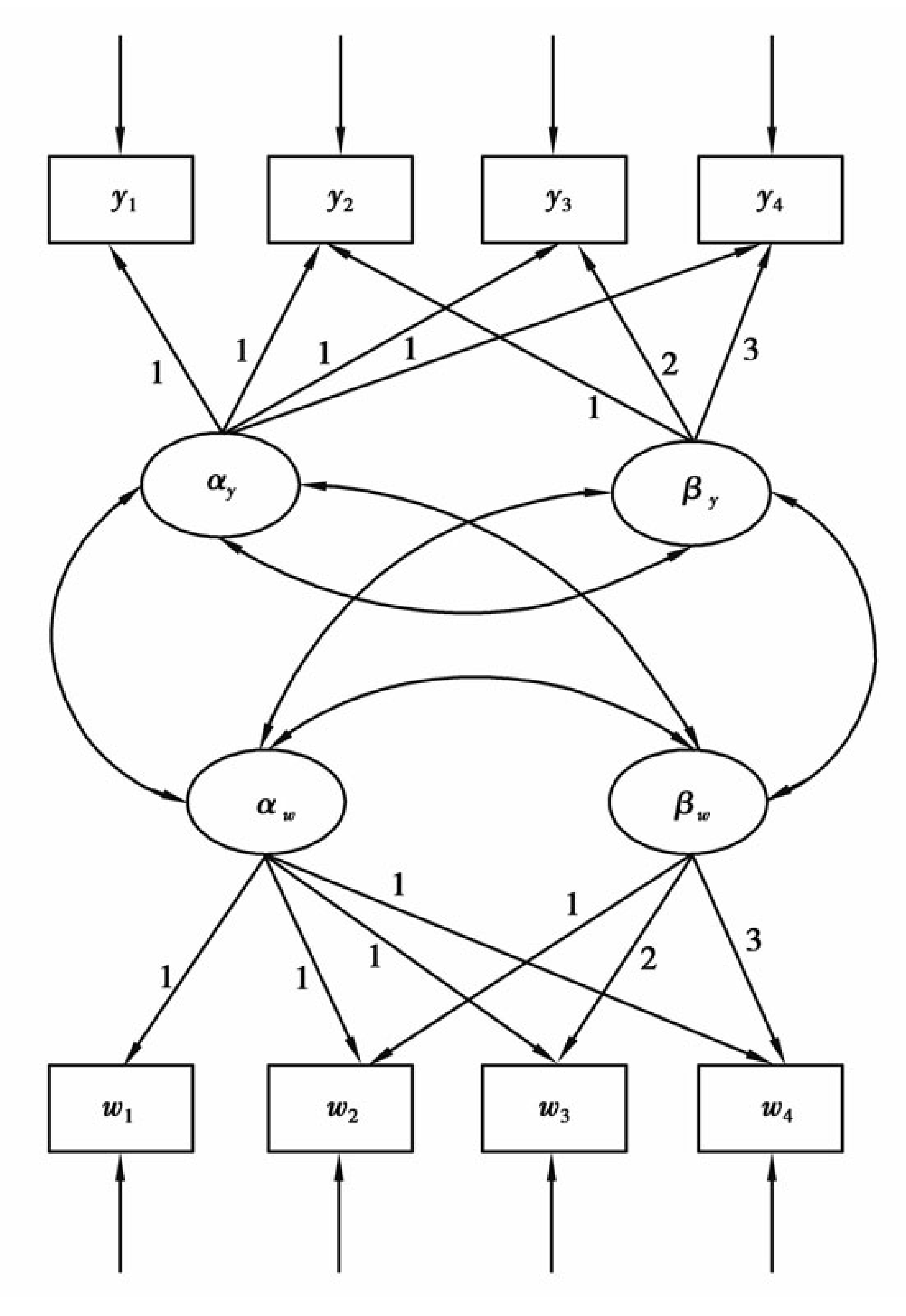


图4-8　多变量增长曲线模型示意图

4.3.4　高阶或多指标LGCM

前面讲述的模型只涉及单个指标重复测量多次，或将多个指标加总后作为单个重复测量数据进行建模的情境。而实际的研究中，多数情况下使用多指标工具重复测量数次，此时获得的数据与单个指标测量数据有些许不同，情况类似于二阶验证性因素分析。具体来说，在重复测量指标上建立一阶因子，然后在一阶因子上建立增长模型，如图4-9，就是一个带有4个测量指标的高阶LGCM。例如，用5个条目的自尊量表每间隔2个月测量同一群体5次，这时如果在指标层面上考察个体自尊发展轨迹，就需要用每次的5个指标定义单个时点的自尊因子，共5个自尊因子，然后将这5个因子作为“指标”建立高阶潜增长模型。

高阶潜增长模型的优点是使用多个指标测量潜变量，由于考虑了测量误差，结果更可靠。当然，高阶潜增长模型涉及测量模型，此时面临测量不变性的问题，即不同时间点测量工具本身是等值的，否则很难解释变化的原因，只有满足测量不变性的前提才能保证增长因子刻画的是因子均值的变化。关于纵向测量不变性问题请参见“基础篇”的第7章（王孟成，2014）。

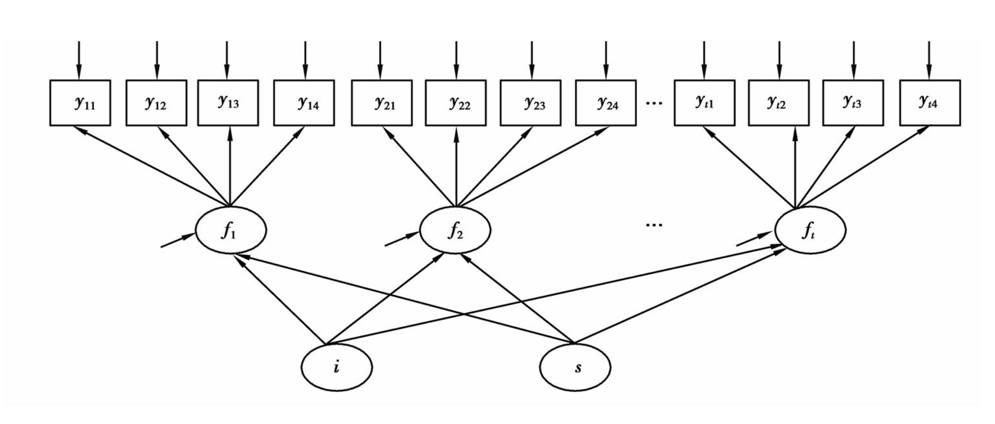


图4-9　高阶LGCM路径图

4.4　增长曲线模型几个重要的问题

4.4.1　测量次数

LGCM最少需要3次测量。主要出于如下考虑：第一，统计识别的角度。LGCM其实就是CFA的一个特殊形式，由于每个测量指标在多个潜变量（如截距和斜率因子）上有负荷，因此需要多次测量。以无条件LGCM为例，两个潜变量至少需要3次测量指标才能识别（3次测量提供6个参数，此时需要估计5个或6个未知参数）。第二，清晰揭示研究问题的角度。在平面上，两点可以确定一条直线，决定直线变化趋势至少需要3个以上的点。第三个点可能存在偶然因素扭曲发展趋势，因此需要第4次测量来确认发展趋势。

4.4.2　样本量

样本量是追踪研究设计的一个重要问题。由于需要对同一批研究对象反复测量多次，研究对象的流失和反感会增加测量难度，通常会采用一些奖励措施，因此样本量都不会太大。但LGCM作为多变量建模方法之一，要满足多变量模型对样本量的一般要求。随着缺失值处理技术的成熟，有计划的缺失研究设计（planned missing data design）可以在不损伤统计功效的前提下增加样本量（Graham et al，2006；Rhemtulla＆Little，2012；王孟成，叶浩生，2014）。

4.4.3　时间分数的确定

不同的时间对中（time centering）或编码（coding）方式影响截距的结果，但不影响斜率的结果[3]。选择对中点取决于研究的目的和结果解释的方便性（Biesanz et al，2004）。如果研究的目的在于揭示增长初始状态时的个体差异、变量间的关系，则选择原始的时间编码（初始测量时间设定为0）；如果研究的兴趣在于最后时点，则选择最后时点对中（最后一次时间点设定为0）。

4.5　M*plus* 具体实例

4.5.1　LGCM在M*plus* 中的设置

增长模型在M*plus* 中有两种不同的定义方法。第一种类似CFA的测量式定义方法，另一种是M*plus* 特有的定义方式。下面简单介绍两种方法的具体设定。

表4-1　增长模型的两种设定方法

	增长式	测量式
截距模型	MODEL： i｜y1-y4；	MODEL： i BY y1-y4； ［y1-y4@0 i］；
线性增长模型	MODEL： i s｜y1@0 y2@1 y3@2 y4@3；	MODEL： i BY y1-y4@1； s By y1@0 y2@1 y3@

		2 y4@3； ［y1-y4@0 i s］；
二次增长模型	MODEL： i s q｜y1@0 y2@1 y3@2 y4@3；	MODEL： i BY y1-y4@1； s By y1@0 y2@1 y3@2 y4@3； q by y1@0 y2@1 y3@4 y4@9； ［y1-y4@0 i s q］；

从表中不难看出，两种方法各有优缺点。增长式的设定简明，出错的机会要少，测量式烦琐但易于理解。至于采用哪种方法全凭个人喜好。

需要注意的是，测量式设定截距因子测量模型时需要设定测量指标的截距为0。因为这里需要估计因子的潜均值，所以采用固定负荷为1，截距为0的方式（对比“基础篇”的第9章）。

符号“｜”左侧为增长因子，右侧为测量时间变量，变量的时间分值通过“@”设定，如果变量的发展趋势是线性的，时间分值之间的间隔是等距的；如果是非线性发展，间隔为非线性。例如，含有4次测量的线性增长模型，第1至第4次测量的时间分值分别设为0，1，2和3。如果是二次增长模型，各时间分值设为0，1，4和9，即是线性增长时的平方。

如果采用增长式设定，M*plus* 包含如下默认设置：①所有指标截距@0；②增长因子均值自由估计；③增长因子协方差自由估计；④指标测量误差彼此不相关。

4.5.2 具体实例说明

下面以具体的实例，分别演示上述模型采用M*plus* 的分析过程，这里主要强调分析过程，而不强调数据结果的理论意义。

中国健康与营养调查（China Health and Nutrition Survey，CHNS）是由美国北卡罗来纳大学和中国疾控中心联合进行的调查。调查从1989年开始，1991、1993、1997、2000、2004、2006、2009、2011年分别进行了追踪调查。该数据已经开放至2009年，可从项目官网下载。CHNS数据样本覆盖大陆9个省份，包含了4400户口26000人。调查内容除了健康营养方面的项目，也涵盖了家庭社会经济等方面的信息，并区分了社区卷（数据需特别申请）和家庭卷。这里选取1989年40～50岁组人群作为分析的对象，关注的问题是这个队列（cohort）平均血压［（收缩压＋2×舒张压）／3］随年龄变化的轨迹。CHNS数据中该队列共有865人，从1989—2006年共7轮数据。由于失访等各种原因，排除掉变量缺失的情况，有效的“人—年数据”（person-year data）共4508个。7次测量血压的均值和相关系数矩阵等描述统计量呈现在表4-2中。

表4-2　7次测量血压的均值和相关系数矩阵

	T89	T91	T93	T97	T00	T04	T06
n	849	761	720	552	571	534	521
均值	88.003	87.837	89.420	92.422	93.988	95.911	96.436
标准差	11.384	12.1786	11.608	13.157	13.133	14.041	14.152
T89	1						
T91	0.527	1					
T93	0.452	0.493	1				
T97	0.467	0.498	0.483	1			
T00	0.357	0.419	0.451	0.556	1		
T04	0.274	0.316	0.296	0.415	0.447	1	
T06	0.319	0.265	0.348	0.406	0.420	0.496	1

注：表中的相关系数均在0.001水平上显著。

4.5.3 无条件线性LGCM示例

（1）语句设定

下面以CHNS的7次血压数据为例，演示M*plus* 的语句设置和结果解释。首先检验CHNS血压数据的无条件LGCM模型，语句和部分输出结果见框4-1。

框4-1 CHNS血压数据的无条件LGCM语句和部分输出结果

```
TITLE: This is an example of unconditional LGCM like figue 4-2
Data: File is chnspress_wide.dat ;
Variable: Names are id gender bmi89 bmi91 bmi93 bmi97 bmi00 bmi04 bmi06
       press89 press91 press93 press97 press00 press04 press06;
      USEVARIABLES=press89 press91 press93 press97 press00 press04 press06;
      Missing are all (-9999) ;
Model:! 采用增长式定义增长模型,也可以采用测量式。
      i s | press89@0 press91@1 press93@2 press97@3 press00@4
      press04@5 press06@6;
Output: tech1 stand;
Plot: Type is Plot3;
       Series=press89 press91 press93 press97 press00 press04 press06(*);
```

```
MODEL FIT INFORMATION

Number of Free Parameters                          12
Loglikelihood
          H0 Value                         -17205.417
          H1 Value                         -17162.257
Information Criteria
          Akaike (AIC)                      34434.834
          Bayesian (BIC)                    34491.986
          Sample-Size Adjusted BIC          34453.877
            (n* = (n + 2) / 24)
Chi-Square Test of Model Fit
          Value                                86.319
          Degrees of Freedom                       23
          P-Value                              0.0000
RMSEA (Root Mean Square Error Of Approximation)
          Estimate                              0.056
          90 Percent C.I.                0.044  0.069
          Probability RMSEA <= 0.05             0.188
CFI/TLI
          CFI                                   0.950
          TLI                                   0.955
Chi-Square Test of Model Fit for the Baseline Model
         Value                               1294.714
          Degrees of Freedom                       21
          P-Value                              0.0000
SRMR (Standardized Root Mean Square Residual)
          Value                                 0.102
MODEL RESULTS
                                                        Two-Tailed
                     Estimate       S.E.    Est./S.E.     P-Value
```

```
I        |
    PRESS89          1.000        0.000     999.000     999.000
    PRESS91          1.000        0.000     999.000     999.000
    PRESS93          1.000        0.000     999.000     999.000
    PRESS97          1.000        0.000     999.000     999.000
    PRESS00          1.000        0.000     999.000     999.000
```

```
    PRESS04          1.000      0.000    999.000    999.000
    PRESS06          1.000      0.000    999.000    999.000
S        |
    PRESS89          0.000      0.000    999.000    999.000
    PRESS91          1.000      0.000    999.000    999.000
    PRESS93          2.000      0.000    999.000    999.000
    PRESS97          3.000      0.000    999.000    999.000
    PRESS00          4.000      0.000    999.000    999.000
    PRESS04          5.000      0.000    999.000    999.000
    PRESS06          6.000      0.000    999.000    999.000
S          WITH
    I               -5.113      1.155     -4.428      0.000
Means ! 增长因子均值
    I               87.236      0.365    239.130      0.000
    S                1.569      0.093     16.870      0.000
Intercepts
    PRESS89          0.000      0.000    999.000    999.000
    PRESS91          0.000      0.000    999.000    999.000
    PRESS93          0.000      0.000    999.000    999.000
    PRESS97          0.000      0.000    999.000    999.000
    PRESS00          0.000      0.000    999.000    999.000
    PRESS04          0.000      0.000    999.000    999.000
    PRESS06          0.000      0.000    999.000    999.000
Variances! 增长因子的方差估计及显著性检验。
    I               77.062      5.645     13.652      0.000
    S                2.546      0.348      7.318      0.000
Residual Variances
    PRESS89         57.719      4.626     12.476      0.000
    PRESS91         74.867      4.817     15.542      0.000
    PRESS93         74.588      4.742     15.729      0.000
    PRESS97         86.191      6.063     14.217      0.000
    PRESS00         89.860      6.417     14.004      0.000
    PRESS04        114.575      8.624     13.286      0.000
    PRESS06        102.447      9.082     11.281      0.000
STANDARDIZED MODEL RESULTS
```

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
I |				
PRESS89	0.756	0.020	38.662	0.000
PRESS91	0.731	0.022	33.989	0.000
PRESS93	0.738	0.024	31.028	0.000
PRESS97	0.704	0.027	26.493	0.000
PRESS00	0.680	0.028	24.691	0.000
PRESS04	0.614	0.027	22.767	0.000
PRESS06	0.606	0.028	21.558	0.000
S |				
PRESS89	0.000	0.000	999.000	999.000
PRESS91	0.133	0.009	14.157	0.000
PRESS93	0.268	0.019	14.003	0.000
PRESS97	0.384	0.027	13.979	0.000
PRESS00	0.494	0.035	14.247	0.000
PRESS04	0.558	0.039	14.291	0.000
PRESS06	0.661	0.044	14.856	0.000
S WITH ！增长因子间相关系数。				
I	-0.365	0.060	-6.111	0.000
Means				
I	9.937	0.365	27.192	0.000
S	0.984	0.091	10.758	0.000
Intercepts				
PRESS89	0.000	0.000	999.000	999.000
PRESS91	0.000	0.000	999.000	999.000
PRESS93	0.000	0.000	999.000	999.000
PRESS97	0.000	0.000	999.000	999.000
PRESS00	0.000	0.000	999.000	999.000
PRESS04	0.000	0.000	999.000	999.000
PRESS06	0.000	0.000	999.000	999.000
Variances				
I	1.000	0.000	999.000	999.000
S	1.000	0.000	999.000	999.000
Residual Variances				
PRESS89	0.428	0.030	14.479	0.000
PRESS91	0.519	0.024	21.849	0.000
PRESS93	0.528	0.022	23.662	0.000
PRESS97	0.554	0.024	22.789	0.000
PRESS00	0.539	0.026	20.846	0.000
PRESS04	0.561	0.028	20.129	0.000
PRESS06	0.488	0.033	14.665	0.000

（2）结果解读

首先检查模型拟合是否达到要求。模型拟合指数分别为：x^2＝86.319，df＝23，p＜0.01，TLI＝0.955，CFI＝0.950，AIC＝34434.834，BIC＝34491.986，RMSEA＝0.056。根据通常的拟合指数评价标准判断，模型拟合良好。

截距因子和斜率因子的方差估计分别为77.062和2.546，均在0.01水平上显著，说明个体在初始水平及增长速度上存在显著个体间差异。截距增长因子和斜率增长因子的相关系数为－0.365（p＜0.01），提示个体增长变化率与初始状态存在显著的反向关系，即初始状态得分高的个体下降速度较慢。

4.5.4 无条件非线性（二次）LGCM示例

（1）语句设定

无条件非线性（二次）LGCM语句和部分结果呈现在框4-2中。二次增长模型的设定非常简单，只需在无条件线性LGCM基础上增加一个二次增长因子即可。M*plus* 会自动设置二次增长模式，即不同时点因子负荷呈2次方增长。

框4-2 无条件非线性（二次）LGCM语句和部分结果

```
TITLE: This is an example of unconditional Quadratic LGCM like figue 4-3
Model:
  i s q | press89@0 press91@1 press93@2 press97@3 press00@4 press04@5 press06@6;
──────────────────────────────────────────────────────────────
MODEL RESULTS
                                                      Two-Tailed
                    Estimate       S.E.   Est./S.E.    P-Value
I          |   ! I 和 S 的输出结果同前
    PRESS89            1.000      0.000    999.000    999.000
S          |
    PRESS89            0.000      0.000    999.000    999.000
Q          | ! Mplus 自动设置二次增长模式即因子负荷呈 2 次方增长。
    PRESS89            0.000      0.000    999.000    999.000
    PRESS91            1.000      0.000    999.000    999.000
    PRESS93            4.000      0.000    999.000    999.000
    PRESS97            9.000      0.000    999.000    999.000
    PRESS00           16.000      0.000    999.000    999.000
    PRESS04           25.000      0.000    999.000    999.000
    PRESS06           36.000      0.000    999.000    999.000
S          WITH
    I                 -8.534      4.533     -1.883      0.060
Q          WITH
    I                  0.495      0.675      0.733      0.463
    S                 -1.831      0.535     -3.420      0.001
Means
    I                 87.574      0.382    229.039      0.000
    S                  1.069      0.253      4.222      0.000
    Q                  0.089      0.044      2.019      0.043
Variances
    I                 79.879      7.807     10.231      0.000
    S                 12.544      3.419      3.669      0.000
    Q                  0.345      0.092      3.762      0.000
STANDARDIZED MODEL RESULTS
STDYX Standardization
                                                      Two-Tailed
                    Estimate       S.E.   Est./S.E.    P-Value
I          |
    PRESS89            0.787      0.032     24.587      0.000
    PRESS91            0.737      0.031     23.891      0.000
    PRESS93            0.741      0.037     20.103      0.000
    PRESS97            0.707      0.037     18.934      0.000
    PRESS00            0.688      0.037     18.457      0.000
    PRESS04            0.623      0.034     18.357      0.000
    PRESS06            0.633      0.035     18.003      0.000
S          |
    PRESS89            0.000      0.000    999.000    999.000
    PRESS91            0.292      0.039      7.446      0.000
    PRESS93            0.587      0.080      7.315      0.000
    PRESS97            0.840      0.115      7.315      0.000
```

PRESS00	1.090	0.151	7.227	0.000
PRESS04	1.234	0.170	7.260	0.000
PRESS06	1.505	0.212	7.111	0.000
Q |				
PRESS89	0.000	0.000	999.000	999.000
PRESS91	0.048	0.006	7.575	0.000
PRESS93	0.195	0.026	7.514	0.000
PRESS97	0.418	0.056	7.489	0.000
PRESS00	0.723	0.098	7.396	0.000
PRESS04	1.024	0.137	7.483	0.000
PRESS06	1.498	0.205	7.322	0.000
S WITH				
I	−0.270	0.107	−2.528	0.011
Q WITH				
I	0.094	0.119	0.794	0.427
S	−0.880	0.031	−28.127	0.000

（2）结果解读

二次增长模型的拟合指数如下：χ^2＝60.087，df＝19，$p<0.01$，CFI＝0.968，TLI＝0.964，AIC＝34416.601，BIC＝34492.805，RMSEA＝0.050。从拟合指数的绝对值上看，二次方增长模型拟合结果优于线性模型，ΔAIC＝18.233，但BIC差异微乎其微。

截距因子与线性斜率因子的相关为－0.27（p＝0.011），二次斜率因子与截距因子和线性斜率因子的相关分别为0.094（p＝0.427）和－0.88（$p<0.001$）。二次斜率因子与截距因子相关不显著，说明初始状态与二次增长因子关系不大。二次增长因子与线性增长因子显著正相关且相关系数较高，提示一次增长率越快的个体二次增长率也越快。

4.5.5 时间分值自由估计的LGCM示例

（1）模型设定

在实际分析过程中，即使设置特定的时间分或增长模式，模型最后估计的结果也可能与之存在较大差异，所以在实践中可以不设定时间分或增长模式，而通过数据直接估计时间分或增长模式，即采用数据驱动的思路。

采用时间分值自由估计的LGCM语句和部分结果呈现在框4-3中。为了模型识别的目的，这里只设定1989年、1991年和1993年的时间分，随后的时间分则由模型自由估计。

框4-3 时间分自由估计的LGCM语句

```
TITLE: This is an example of a LGCM with free time score in figue 4-4
Model:
        i s | press89@0 press91@1 press93@2 press97 press00 press04 press06;
————————————————————————————————————————————

MODEL RESULTS
                                                         Two-Tailed
                     Estimate       S.E.    Est./S.E.     P-Value
 I          !! 同前
    PRESS89             1.000      0.000    999.000     999.000
 S          !! 从第 4 次测量开始,增长模式由软件通过数据自由估计
    PRESS89             0.000      0.000    999.000     999.000
    PRESS91             1.000      0.000    999.000     999.000
    PRESS93             2.000      0.000    999.000     999.000
    PRESS97             6.146      1.452      4.232       0.000
    PRESS00             8.607      2.048      4.202       0.000
    PRESS04            11.695      2.877      4.065       0.000
    PRESS06            11.868      2.928      4.054       0.000
S          WITH
    I                  -1.778      0.746     -2.383       0.017
Means
    I                  87.748      0.378    232.427       0.000
    S                   0.722      0.189      3.826       0.000
```

（2）结果解读

模型拟合指标：χ^2 =53.212，df =19，p <0.01，CFI＝0.973，TLI＝0.970，AIC＝34409.726，BIC＝34485.930，RMSEA＝0.046。与固定时间分的LGCM相比，自由估计的增长模型的拟合有了显著提升[(4)]，ΔAIC＝25.108，ΔBIC＝6.056。此时，截距因子和斜率因子的方差估计分别为73.071（p <0.001）和0.50；斜率因子的方差变得不再显著（p =0.063），说明个体间增长趋势的差异无统计学意义。截距因子与斜率因子的相关系数也由−0.365降为−0.294（p <0.001）。尽管存在上述差异，两种模型整体上仍表现出类似的发展趋势。

4.5.6　包含时间变化和不变化协变量的LGCM示例

（1）模型设定

这里我们设定2种类型的协变量：时间变化和不变化协变量。其中时间变化的协变量为与血压同时采集的BMI数据，时间不变的协变量采用被试性别。语句和部分结果呈现在框4-4中。

框4-4　包含时间变化和不变化协变量的LGCM语句

```
TITLE: This is an example of a LGCM with time-varying and invariant covariates
Data:  File is chnspress_wide.dat;
Variable: Names are id gender bmi89 bmi91 bmi93 bmi97 bmi00 bmi04 bmi06
       press89 press91 press93 press97 press00 press04 press06;
  Missing are all (-9999) ;
USEVARIABLES=gender press89 press91 press93 press97 press00 press04 press06 bmi89 bmi91 bmi93
bmi97 bmi00 bmi04 bmi06;
Model:
    i s | press89@0 press91@1 press93@3 press97 press00 press04 press06;
    i s on gender;! 定义图 4-5 中的协变量影响增长因子的路径
    press89 on bmi89;! 下面几行定义图 4-6 中含时间变化协变量的增长模型。
    press91 on bmi91;
    press93 on bmi93;
    press97 on bmi97;
    press00 on bmi00;
    press04 on bmi04;
    press06 on bmi06;
```

```
  Output: sampstat tech1 stand;
______________________________________________________________
STANDARDIZED MODEL RESULTS

I          ON   ! 不变协变量性别的路径系数。
    GENDER              0.296        0.065        4.529        0.000
S          ON
    GENDER              0.044        0.090        0.490        0.624
PRESS89   ON! 时间变化协变量路径系数。
    BMI89               0.308        0.038        8.145        0.000
PRESS91   ON
    BMI91               0.302        0.036        8.497        0.000
PRESS93   ON
    BMI93               0.323        0.037        8.788        0.000
PRESS97   ON
    BMI97               0.338        0.036        9.340        0.000
PRESS00   ON
    BMI00               0.324        0.037        8.832        0.000
PRESS04   ON
    BMI04               0.271        0.043        6.322        0.000
PRESS06   ON
    BMI06               0.216        0.054        4.001        0.000
S          WITH
    I                  -0.253        0.116       -2.176        0.030
```

（2）结果解读

模型拟合指标：$\chi^2=116.322$，$df=66$，$p<0.01$，CFI＝0.936，TLI＝0.926，AIC＝15393.08，BIC＝15484.74，RMSEA＝0.051，结果提示模型拟合尚可。

协变量的预测效果表明，性别对斜率和截距因子的回归系数分别为0.044（$p=0.624$）和0.296（$p<0.001$），说明不同性别参与者在血压的初始水平存在显著性别差异，而在血压变化斜率上差异不显著。同时BMI均对同时期的血压产生显著的预测作用，回归系数在0.30上下（$ps<0.001$）。

由于纳入了时间变化和时间不变的协变量（性别和BMI），此时模型的结果为控制上述变量后的结果。如果研究的兴趣是BMI对血压的影响，此时关注的侧重点应该是BMI对血压的回归系数。如果兴趣仍然是血压变化的趋势，此时侧重点同无条件模型，只是此时的模型考虑了协变量，为控制相关协变量后的结果。

4.5.7　平行发展模式的LGCM示例

（1）模型设定

前面我们提到，将同时采集的协变量按4.5.6部分的建模方式处理存在忽略测量误差等不足，而将其作为平行的发展过程可以更好地避免上述不足。框4-5的语句将BMI数据作为独立的LGCM过程分析，通过设置增长因子间的回归方程考查两者之间的影响过程[5]。

框4-5　平行发展模式的LGCM语句

```
TITLE: This is an example of a LGCM with two parallel processes
Data:  File is chnspress_wide.dat ;
USEVARIABLES=press89 press91 press93 press97 press00 press04 press06
  bmi89 bmi91 bmi93 bmi97 bmi00 bmi04 bmi06;
Analysis:  processor=2;
Model:! 定义2个独立的增长模型
     ipress spress | press89@0 press91@1 press93@2 press97@3 press00 press04 press06;
     ibmi sbmi | bmi89@0 bmi91@1 bmi93@3 bmi97 bmi00 bmi04 bmi06;
     spress ipress on sbmi ibmi;
Output: sampstat tech1 stand;
——————————————————————————————————————————————————

MODEL RESULTS
                                                        Two-Tailed
                    Estimate       S.E.    Est./S.E.     P-Value
SPRESS   |
    PRESS89            0.000      0.000      999.000     999.000
    PRESS91            1.000      0.000      999.000     999.000
    PRESS93            2.000      0.000      999.000     999.000
    PRESS97            3.000      0.000      999.000     999.000
    PRESS00            5.060      0.458       11.054       0.000
    PRESS04            6.740      0.642       10.505       0.000
    PRESS06            6.772      0.652       10.384       0.000
  SBMI     |
    BMI89              0.000      0.000      999.000     999.000
    BMI91              1.000      0.000      999.000     999.000
    BMI93              3.000      0.000      999.000     999.000
    BMI97              6.703      0.692        9.680       0.000
    BMI00              9.274      0.971        9.548       0.000
    BMI04             12.528      1.310        9.564       0.000
    BMI06             12.079      1.284        9.405       0.000
SPRESS   ON
    SBMI               3.287      0.750        4.381       0.000
    IBMI              -0.063      0.030       -2.135       0.033
```

```
IPRESS   ON
    SBMI              -1.636      2.878       -0.568       0.570
    IBMI               1.393      0.129       10.836       0.000
SPRESS   WITH
    IPRESS            -3.148      0.933       -3.374       0.001
SBMI     WITH
    IBMI               0.003      0.020        0.157       0.875
Means
    IBMI              22.100      0.095      232.381       0.000
    SBMI               0.103      0.013        7.730       0.000
Intercepts
    IPRESS            56.794      2.852       19.910       0.000
    SPRESS             2.351      0.669        3.514       0.000
Variances
    IBMI               7.229      0.371       19.487       0.000
    SBMI               0.024      0.006        4.353       0.000
STANDARDIZED MODEL RESULTS
STDYX Standardization
```

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
IPRESS |				
PRESS89	0.747	0.019	39.471	0.000
PRESS91	0.727	0.021	34.993	0.000
PRESS93	0.734	0.022	32.907	0.000
PRESS97	0.698	0.025	27.844	0.000
PRESS00	0.676	0.027	24.729	0.000
PRESS04	0.606	0.027	22.451	0.000
PRESS06	0.600	0.027	22.384	0.000
SPRESS |				
PRESS89	0.000	0.000	999.000	999.000
PRESS91	0.109	0.013	8.571	0.000
PRESS93	0.219	0.026	8.591	0.000
PRESS97	0.313	0.037	8.437	0.000
PRESS00	0.511	0.042	12.140	0.000
PRESS04	0.610	0.046	13.186	0.000
PRESS06	0.607	0.045	13.484	0.000
IBMI |				
BMI89	0.937	0.005	178.432	0.000
BMI91	0.959	0.005	184.658	0.000
BMI93	0.922	0.009	102.067	0.000
BMI97	0.852	0.015	56.427	0.000
BMI00	0.797	0.018	45.270	0.000
BMI04	0.765	0.020	37.555	0.000
BMI06	0.765	0.020	38.353	0.000

SBMI |				
BMI89	0.000	0.000	999.000	999.000
BMI91	0.056	0.007	8.508	0.000
BMI93	0.160	0.018	8.688	0.000
BMI97	0.331	0.021	15.902	0.000
BMI00	0.429	0.022	19.752	0.000
BMI04	0.556	0.023	23.780	0.000
BMI06	0.536	0.024	22.533	0.000
SPRESS ON				
SBMI	0.395	0.065	6.067	0.000
IBMI	−0.130	0.060	−2.162	0.031
IPRESS ON				
SBMI	−0.029	0.051	−0.573	0.566
IBMI	0.431	0.037	11.794	0.000
SPRESS WITH				
IPRESS	−0.340	0.068	−4.971	0.000
SBMI WITH				
IBMI	0.008	0.048	0.157	0.875
Means				
IBMI	8.220	0.214	38.469	0.000
SBMI	0.657	0.052	12.680	0.000

（2）结果解读

模型拟合指标：x^2 ＝297.467，df ＝84，p ＜0.01，CFI＝0.97，TLI＝0.967，AIC＝51523.273，BIC＝51689.968，RMSEA＝0.054，结果提示模型拟合很好。

BMI的截距因子与斜率因子间的相关系数为0.008（p ＝0.875），提示BMI的初始水平与随后的发展速率无关。BMI截距因子到血压截距和斜率因子的回归系数分别为0.431（p ＜0.01）和－0.13（p ＝0.031），说明BMI的初始水平对血压的初始水平有显著的正向预测作用，而对血压的变化速率有显著的负向预测作用。

BMI斜率因子到血压截距和斜率因子的回归系数分别为－0.029（p ＝0.566）和0.395（p ＜0.01），结果说明BMI的变量速率对血压的初始水平没有显著的预测作用，而对血压的变化速率有显著的正向预测作用。上述结果与理论预期基本一致。图4-10呈现了平行LGCM的路径。

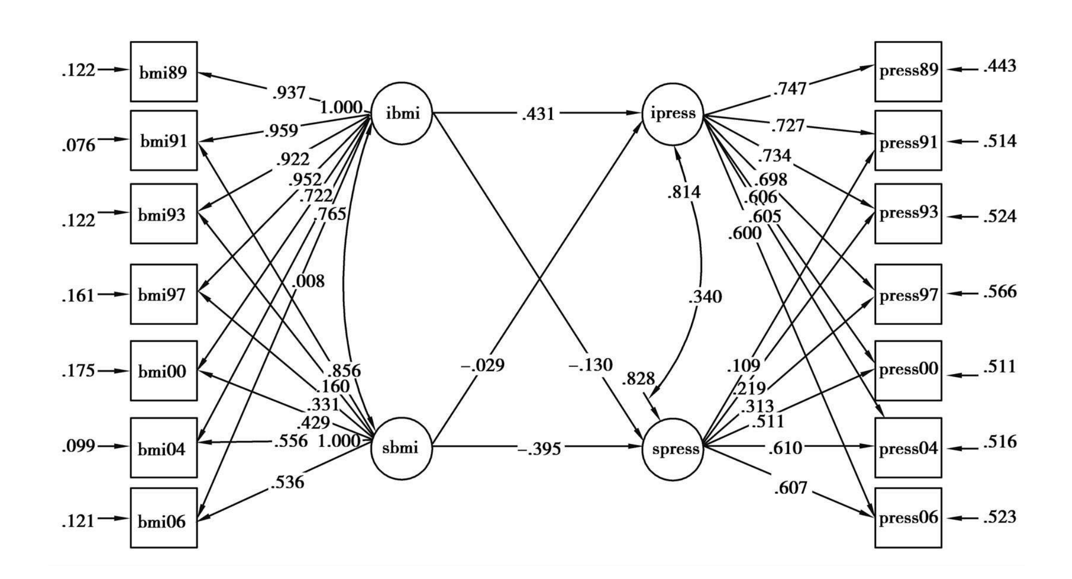


图4-10　平行LGCM路径图

4.5.8　多组增长模型

（1）模型设定

有时研究者会对不同组别的个体在某种心理或行为上的发展趋势的差异感兴趣。比如，上述血压的例子，我们可以比较不同性别个体的发展趋势是否存在差异。当然，这种检验也可以通过引入性别协变量来进行，但采用分组的方法可以比较更多的参数是否等值。下面我们在框4-5的基础上给出一个简单的多组LGCM的示例，更多的应用大家可以结合相关理论和M*plus* 操作的知识进行组合。

框4-6　多组LGCM语句示例

```
TITLE: This is an example of a LGCM with two parallel processes
Data: File is chnspress_wide.dat ;
Variable: Names are id gender bmi89 bmi91 bmi93 bmi97 bmi00 bmi04 bmi06
      press89 press91 press93 press97 press00 press04 press06;
  Missing are all (-9999) ;
USEVARIABLES=press89 press91 press93 press97 press00 press04 press06
  bmi89 bmi91 bmi93 bmi97 bmi00 bmi04 bmi06;
  grouping=gender(1=male 0=female);
Analysis: processor=2;
Model: ! 定义 2 个独立的增长模型。
 ipress spress | press89@0 press91@1 press93@2 press97 press00 press04 press06;
 ibmi sbmi | bmi89@0 bmi91@1 bmi93@2 bmi97 bmi00 bmi04 bmi06;
      spress on sbmi ibmi;
      ipress on sbmi ibmi;
Model female:
```

```
 ipress spress | press89@0 press91@1 press93@2 press97 press00 press04 press06;
  ibmi sbmi | bmi89@0 bmi91@1 bmi93@2 bmi97 bmi00 bmi04 bmi06;
      spress on sbmi ibmi;
      ipress on sbmi ibmi;
Output: sampstat tech1 stand;
```

```
STANDARDIZED MODEL RESULTS
                                                        Two-Tailed
                      Estimate       S.E.    Est./S.E.     P-Value
Group FEMALE
    IPRESS    |
    PRESS89              0.789      0.023       33.904       0.000
    PRESS91              0.742      0.025       29.371       0.000
    PRESS93              0.760      0.026       28.708       0.000
    PRESS97              0.743      0.035       21.445       0.000
    PRESS00              0.724      0.036       19.997       0.000
    PRESS04              0.680      0.039       17.469       0.000
    PRESS06              0.693      0.040       17.399       0.000
SPRESS    ON
    SBMI                 0.404      0.091        4.435       0.000
    IBMI                -0.070      0.085       -0.823       0.410
IPRESS    ON
    SBMI                -0.066      0.069       -0.955       0.339
    IBMI                 0.410      0.050        8.251       0.000
SPRESS    WITH
    IPRESS              -0.387      0.090       -4.324       0.000
SBMI      WITH
    IBMI                 0.015      0.066        0.224       0.823
Means
    IBMI                 7.994      0.291       27.460       0.000
    SBMI                 0.709      0.074        9.634       0.000
```

Group MALE

```
  IPRESS    |
    PRESS89              0.648      0.030       21.623       0.000
```

PRESS91	0.670	0.032	20.729	0.000
PRESS93	0.660	0.032	20.605	0.000
PRESS97	0.601	0.036	16.601	0.000
PRESS00	0.574	0.037	15.457	0.000
PRESS04	0.485	0.033	14.490	0.000
PRESS06	0.466	0.033	14.307	0.000
SPRESS ON				
SBMI	0.391	0.094	4.175	0.000
IBMI	−0.223	0.086	−2.583	0.010
IPRESS ON				
SBMI	0.089	0.076	1.174	0.240
IBMI	0.540	0.051	10.537	0.000
SPRESS WITH				
IPRESS	−0.140	0.132	−1.057	0.291
SBMI WITH				
IBMI	−0.037	0.070	−0.522	0.601
Means				
IBMI	8.719	0.327	26.637	0.000
SBMI	0.616	0.073	8.429	0.000

（2）结果解读

模型拟合指标：x^2 ＝373.367，df ＝166，p ＜0.01，CFI＝0.970，TLI＝0.967，AIC＝51426.121，BIC＝51769.037，RMSEA＝0.054，结果提示模型拟合很好。

框4-6还呈现了男女两组的结果，从中不难发现，两组负荷和回归系数存在差异。如果想检验某条回归系数差异是否显著，可以使用model test命令实现，具体参照“基础篇”的相关内容。

4.6 本章小结

增长曲线模型是揭示事物发展进度最常用的建模方法。该方法具有非常灵活的建模策略，本章主要介绍了最基本的建模形式，其他更复杂的模型类型可以在上述基本模型的基础上进行修改。例如可以在平行发展LGCM中加入中介变量和协变量。另外，LGCM并不是刻画事物增长的唯一方法，多水平模型和考虑样本异质性的增长混合模型亦是备选方法，后面的章节将对它们进行介绍。

※推荐阅读※

LGCM已不算新技术，早在10多年前就有不少介绍文章。如果读者想要系统学习LGCM方法，可以参考Bollen＆Curran（2006）的专著。

Bollen，K．A．，＆Curran，P．J．（2006）．*Latent Curve Models：A Structural Equation Perspective*． Hoboken，NJ：Wiley.

---

(1) 用于多水平建模的软件也有很多，常用的有HLM（Raudenbush，Bryk＆Cheong，2008），MLwiN（Rasbash，Steele，Browne＆Goldstein，2009），SPSS近年也新增了多水平模块。

(2) 主要出于如下几点考虑：首先，潜在（Latent），说明是潜变量模型；其次，增长（Growth），表明分析的是纵向数据；最后，曲线（Curve），说明发展的轨迹往往是曲线的。另外，LCM和LTM很容易和潜类别模型（Latent Class Modeling）和潜在转换模型（Latent Transition Modeling）混淆。

(3) 对于二次增长或多次增长模型，时间编码不影响最高次增长因子（均值和方差）的估计。例如，时间编码只影响截距和一次斜率的估计结果，二次增长因子的结果不受编码方式的影响。

(4)　鉴于时间分自由估计模型比固定时间分模型拟合显著提升，随后的分析均基于此模型。

(5)　也可以设置成相关，具体类型主要取决于研究假设。

5　增长混合模型——潜类别增长模型与增长混合模型

在本书的开篇和第4章分别介绍了探测群体异质性（heterogeneity）的潜类别分析（LCA）和描述个体特征随时间变化的潜增长曲线模型（LGCM）。第3章介绍了将LCA与传统的因子分析模型相结合形成的因子混合模型（FMM），用于解释因子分析时群体异质性的问题。与之类似，在LGCM分析中，群体也可能存在异质性，因此可以将两者结合用以解释不同潜在类别组个体增长变化的差异和特点，这种模型称作增长混合模型（Growth Mixture Modeling，GMM；Muthén＆Shedden，1999；王孟成，毕向阳，叶浩生，2014）。

近年来，GMM广泛用于各学科领域，例如心理学（Koss，George，Davies，et al，2013）、教育学（Parrila et al，2005）、公共卫生（Abroms et al，2005）、管理学（Qureshi＆Fang，2011）和犯罪学（Morris＆Slocum，2012）。不仅用于问卷调查类数据的分析，也用于生物学、生理学和临床实验数据的分析（e．g．，Hunter et al，2010；Muthén＆Brown，2009）。

在GMM发展过程中还存在另外一种模型即所谓的基于群组的增长模型（Group-based Trajectory Modeling；Nagin，1999；2005）或潜类别增长模型（Latent Class Growth Analysis，LCGA）。LCGA与GMM的区别主要存在于类别组内的发展轨迹是否允许存在个体变异，因此可以将LCGA作为GMM的特例。

本章的内容安排如下：首先介绍LCGA和GMM两个模型的特点和差异，然后介绍LCGA和GMM的建模过程。最后，通过一个实例简要演示两类模型在M*plus* 中的设置、分析过程和结果解释。

5.1　混合增长模型概述

在传统的增长曲线模型（LGCM）中假设群体同质（homogeneous），群体内所有个体享有相同或类似的平均增长轨迹（即相同的截距和斜率）。

在有些情况下，个体间的差异较大，并不一定能满足LGCM关于样本来自同一群体的假设前提。如图5-1所示，尽管大部分个体血压的增长服从“平稳”的发展的轨迹，但仍有部分个体呈现“先高后缓慢下降”和“先低后加速增长”的发展轨迹[(1)]。当研究的样本呈现如图5-1所示的发展轨迹时，传统增长模型的群体同质性假设受到严重威胁，因为并非所有个体遵从相同的增长趋势。这个时候，传统的增长曲线模型将不再适合，合理的增长模型应该考虑到群体的异质性。

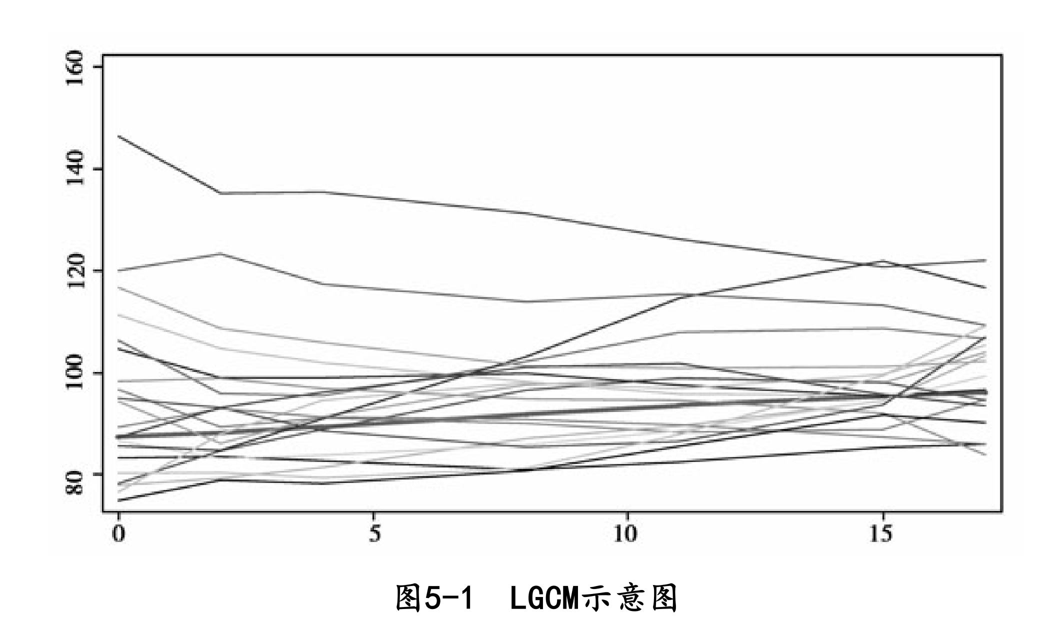


图5-1　LGCM示意图

群体异质性的增长模型有GMM和基于群体的增长模型或LCGM两种。与LGCM不同，GMM和LCGM并不假设群体同质而是允许群体内存在异质性。因此，在GMM和LCGM中同时存在两种潜变量：（1）连续潜变量，用于描述初始差异和发展趋势的随机截距和随机斜率因子，同LGCM；

（2）类别潜变量，通过将群体分成互斥的潜类别亚组来描述群体的异质性。GMM和LCGM如图5-2所示。

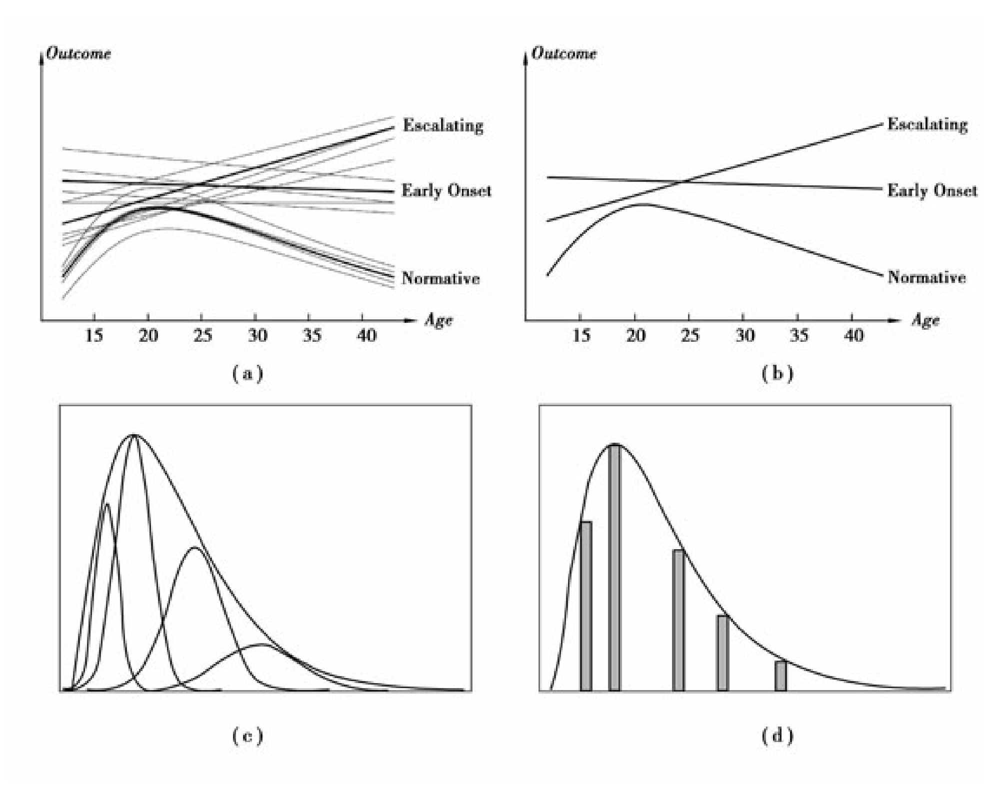


图5-2　GMM和LCGM示意图

图5-2呈现了一例有关酗酒行为研究中的GMM和LCGM的分布形态及增长模式。具体来说，（a）和（b）分别表示GMM和LCGM的增长模式图（Muthén，2008），（c）和（d）分别对应某个时点各自分布的截面图（Feldman，Masyn＆Conger，2009）。从图中可以非常直观地看出两个模型间的区别。（a）和（b）中的粗线代表类别组内个体平均的增长曲线，同LGCM。（a）图中除了粗线外还有代表类别组内个体差异的增长曲线（细线），而（b）图没有细线表示类别组内不存在个体差异（类别组内方差为0）。

另外，可以从多组分析的角度去理解GMM和LCGM。例如，在青少年危险行为研究中发现，男女生在多个类别危险行为发生率上存在差

异，如果要考察男女生吸烟行为随年龄发展变化趋势时则需要将男女生分开，分别建立增长模型。与之类似，GMM和LCGM可以看作多组LGCM，只不过分组的依据是不可观测的潜在类别组（同FMM）。再者，多组分析是事先根据某一分类变量（如性别、民族）将样本划入不同的组别，然后以每个组的样本为对象进行建模（通常是同时进行）。而潜在类别组的划分是基于模型估计的，是潜在的、事后分组的。

表5-1　混合模型与多组分析比较

分组依据 \ 模型	回归分析	因子分析	潜增长模型	结构方程模型
观测类别:多组模型	分组回归	多组因子分析	分组潜增长模型	多组结构方程模型
潜在类别:混合模型	回归混合模型	因子混合模型	增长混合模型	结构方程混合模型

5.2　增长混合模型的表达式

增长混合模型可以理解为，先将群体分成若干潜类别组，然后在每个类别组内建立LGCM用于描述类别组内个体随时间变化的情况。不同的是，在GMM和LCGM中这两步是同时完成的，描述个体随时间变化的测量指标同时也是潜类别分析的指标。GMM和LCGM的方程表达式与LGCM类似，只是将其扩展到包含潜类别变量的形式，如下：

LGCM：

$$
\begin{aligned}
y_{it} &= \alpha_{it} + \beta_{it}\lambda_{it} + \varepsilon_{it} \\
\alpha_i &= \mu_\alpha + \zeta_{\alpha i} \\
\beta_i &= \mu_\beta + \zeta_{\beta i}
\end{aligned}
\qquad (5.1)
$$

GMM：

$$
\begin{aligned}
&y_{it} = \sum_{k=1}^{k} p(c_i = k)[a_{itk} + \lambda_t \beta_{itk} + \varepsilon_{yitk}] \\
&\alpha_{ki} = \mu_{\alpha k} + \zeta_{\alpha ik} \\
&\beta_{ki} = \mu_{\beta k} + \zeta_{\beta ik}
\end{aligned}
\tag{5.2}
$$

LCGM：

$$
\begin{aligned}
&y_{it} = \sum_{k=1}^{k} p(c_i = k)[a_{itk} + \lambda_t \beta_{itk} + \varepsilon_{yitk}] \\
&\alpha_{ki} = \mu_{\alpha k} \\
&\beta_{ki} = \mu_{\beta k}
\end{aligned}
\tag{5.3}
$$

C 为类别潜变量，共有k 个水平，p 为类别概率，i 为个体，t 为测量时间。μ_α 和μ_β 分别表示全部个体截距和斜率的均值即总均值；$\zeta_{\alpha i}$ 和$\zeta_{\beta i}$ 分别表示个体截距和斜率与对应的总均值间的差异，每个个体均有一个特定的值。$\mu_{\alpha k}$、$\mu_{\beta k}$ 和$\zeta_{\beta ik}$ 表示类别特定的相应参数。

根据上述模型方程表达式及示意图不难发现三者之间的关系：LGCM和LCGM是GMM的特例。LGCM假设所有个体来自同质性群体，所有个体享有相同的平均增长曲线，个体与平均增长曲线的差异（个体间差异）通过增长因子方差来表达。当潜类别变量只有一个水平，即$k=1$时，GMM简化为LGCM。

当类别组内不存在方差变异即将类别组内的增长因子方差固定为0时，单个类别组内所有个体具有完全相同的平均增长曲线，此时称为LCGM。GMM则兼而有之，估计类别组内的平均增长曲线的同时也估计增长因子的方差。

5.3　GMM建模的过程

5.3.1　建模过程

总的来说，GMM作为潜变量模型中的一员其建模过程遵循一般潜变量建模的通用流程，但在一些细节上存在不同。下面简要介绍GMM建模过程，与其他潜变量模型建模过程相同或类似之处将不展开讨论。

（1）研究假设

选择GMM作为分析方法的依据。可以从数据和理论假设两个方面来论证选择GMM的依据。当不考虑群体异质性的LGCM模型拟合不理想时，尝试使用GMM来拟合数据，此时为探索性分析。如果根据现有理论假设，可以确定研究群体存在明确的群体异质性，此时使用GMM具有验证性的特点。

（2）模型设定

同LCA一样，GMM模型的设定从基线模型开始。单类别增长模型，然后逐渐增加类别个数。理想的情况是，研究根据理论预测设定潜类别数目，但实际应用上多采用探索的思路，根据数据拟合结果选择模型。

（3）模型估计与评价

GMM的估计方法依据观测指标的类型而定。传统上，GMM的指标为观测变量，但在更宽泛的意义上来说（M*plus* 下的一般潜变量模型框架），GMM可以构建任何尺度变量的模型[(2)]。当观测变量为连续变量时，M*plus* 默认的参数估计方法为稳健较大似然估计（MLR）。除此之外，M*plus* 还支持贝叶斯估计，该方法在最近的模拟研究中表现不俗（Sarah，2013）。

模型评价遵循两条原则：实际意义和统计指标。实际意义的评价应结合理论、前人的研究结果和结果的可理解性。统计指标的讨论见第1章LCA的部分内容。GMM作为潜类别分析模型的一种，遵循着LCA模型选择的策略。

（4）模型结果解释

在确定潜类别个数后，需要对模型的理论意义进行解释，或者引入协变量对不同类别增长趋势进行预测。

5.3.2　包含协变量的GMM

研究实践中除了使用增长模型刻画个体发展过程外，还需要考察其他关键变量对此过程产生的影响或受此过程影响的结局变量。图5-3是一个包含协变量和结局变量的GMM示意图。图中x 为自变量，可以是

任何类型（连续、类别或计数型等）的变量，由x 指向i 和s 的路径表示线性回归，x 指向c 的路径代表多项式回归。潜类别变量c 指向结局变量u 的路径依据结局变量的类型而定。

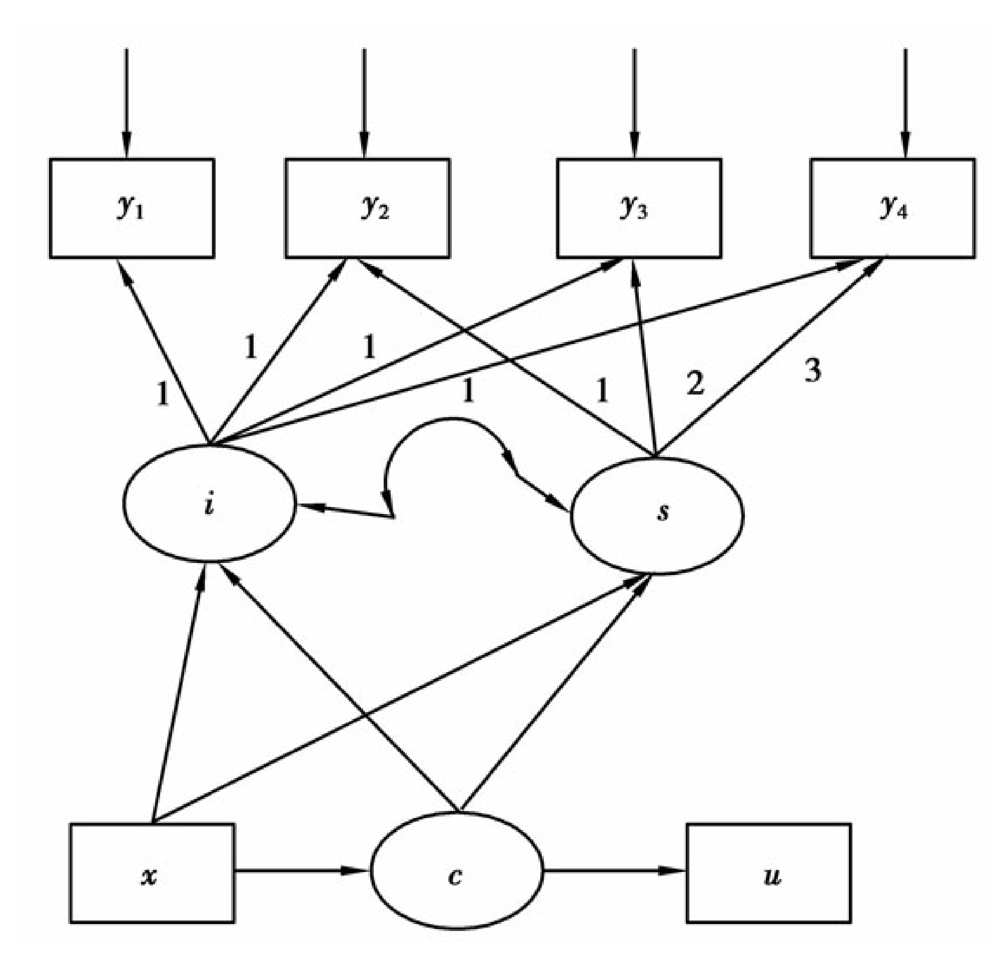


图5-3　包含协变量和结局变量的GMM路径图（来源：M*plus* 手册）

研究发现，包含协变量不仅对探讨变量之间的关系有重要意义，对GMM的潜类别个数的确定也同样重要（Muthén，2004）。然而最近的模拟研究结果并不一致（Li & Hser，2011；Tofighi & Enders，2007），所以现在下结论为时尚早。

5.3.3　样本量

影响GMM估计准确性的因素有测量次数（指标个数）、样本量、类别间距（类别组间的距离大小）以及数据缺失的比例。但目前的研究非常稀少，一方面是GMM刚刚兴起，另一方面应用者学习、掌握应用还有个时间差。但就目前为数不多的模拟研究来看，样本量的要求与上述4个变量有关。在其他条件不变的情况下，样本量随测量次数增加而降低，当到一定程度时影响趋弱。

具体来说，Kim（2012）的模拟研究发现，在类别个数只有2个，类别距离1.5和3个标准差单位，不存在缺失，只有4次测量（单指标）时，最低样本量要求分别为700和200。当类别个数增加到6个时，样本量要求分别增加至2400和300。Kim的研究还发现，在高类别距离的条件下，测量次数的影响有限，但在类别距离较小时，测量次数的影响变大。通常一个研究中，类别间距和数据缺失都很难控制，样本量和测量次数相对比较容易实现，所以在样本量不能增加的情况下，可以通过增加测量次数来增加结果估计的精确性。

5.3.4 GMM在M*plus* 中的设定

M*plus* 在分析GMM时包含如下默认设置：①所有测量指标截距固定为0；②增长因子均值自由估计，并且跨类别变化；③增长因子方差和协方差自由估计，同时限定跨类别等同；④观测指标的残差方差自由估计并且彼此不相关（见Muthén & Muthén，1998—2015，p213-216）。如果设定增长因子方差和协方差为0，模型则为LCGM。

另外，当测量指标为类别变量时，M*plus* 的设置有所变化。具体来说，当“ANALYSIS”后面没有加“ALGORITHM＝INTEGRATION；”

时，M*plus* 执行的是LCGA，当加上“ALGORITHM＝INTEGRATION；”时执行的是GMM（见Muthén＆Muthén，1998—2015，p219）。

如上文所述，我们可以从多组分析的角度去理解GMM，同多组分析一样，也存在等值的问题。除了上述M*plus* 默认的参数设置外，还可以根据研究需要设定模型中的参数等值或释放默认设置允许自由估计。下面先按照M*plus* 的默认设置分别给出GMM和LCGM的语句，后面简要说明其他等值设定的问题。

下面继续使用中国健康与营养调查的数据演示GMM和LCGM的建模过程。

5.4　LCGM分析示例

依据混合模型建模的一般步骤，依次提取了1～6个类别的LCGM模型[3]，拟合结果呈现在表5-2中，取3个类别时的M*plus* 语句呈现在框5-1中。

框5-1　3个类别时LCGM的语句

```
Title: an example of Latent Class Growth Modeling
Data: File is chnspress_wide.dat ;
Variable: Names are id gender bmi89 bmi91 bmi93 bmi97 bmi00 bmi04 bmi06
          press89 press91 press93 press97 press00 press04 press06;
      Missing are all (-9999) ;
USEVARIABLES= press89 press91 press93 press97 press00 press04 press06;
            Classes=c(3);
Analysis: Type=Mixture ;
            Starts=200 10;
            Processor=2;
            LRTBOOTSTRAP=200; ! 获得更准确的LRT估计
Model:
    %Overall%
    i s | press89@0 press91@1 press93@3 press97* press00* press04* press06* ;
      i@0 s@0; ! 设定增长因子方差为0,即为LCGM
Output: tech11 tech14;
Plot: Type is Plot3;
      Series=press89 press91 press93 press97 press00 press04 press06(*);
      Savedata: SAVE=CPROB;
            FILE IS presslgcm.txt;
------------------------------------------------------------------

MODEL RESULTS

                                                      Two-Tailed
                     Estimate       S.E.    Est./S.E.    P-Value

Latent Class 1
```

```
I        |
    PRESS89          1.000       0.000     999.000     999.000
    PRESS91          1.000       0.000     999.000     999.000
    PRESS93          1.000       0.000     999.000     999.000
    PRESS97          1.000       0.000     999.000     999.000
    PRESS00          1.000       0.000     999.000     999.000
    PRESS04          1.000       0.000     999.000     999.000
    PRESS06          1.000       0.000     999.000     999.000
S        |
    PRESS89          0.000       0.000     999.000     999.000
    PRESS91          1.000       0.000     999.000     999.000
    PRESS93          3.000       0.000     999.000     999.000
    PRESS97          7.630       1.758       4.340       0.000
    PRESS00         11.653       2.560       4.552       0.000
    PRESS04         14.896       3.424       4.351       0.000
    PRESS06         14.502       3.310       4.382       0.000
Means! 均值自由估计
```

Means！均值自由估计				
I	114.575	2.803	40.882	0.000
S	−0.308	0.278	−1.107	0.268
Intercepts				
PRESS89	0.000	0.000	999.000	999.000
PRESS91	0.000	0.000	999.000	999.000
PRESS93	0.000	0.000	999.000	999.000
PRESS97	0.000	0.000	999.000	999.000
PRESS00	0.000	0.000	999.000	999.000
PRESS04	0.000	0.000	999.000	999.000
PRESS06	0.000	0.000	999.000	999.000
Variances！方差限定为 0				
I	0.000	0.000	999.000	999.000
S	0.000	0.000	999.000	999.000
Residual Variances				
PRESS89	72.893	5.116	14.249	0.000
PRESS91	83.050	6.552	12.675	0.000
PRESS93	84.123	6.207	13.553	0.000
PRESS97	96.533	7.867	12.271	0.000
PRESS00	108.303	8.846	12.243	0.000
PRESS04	150.587	15.086	9.982	0.000
PRESS06	150.281	18.378	8.177	0.000

```
Latent Class 2
I        |
    PRESS89            1.000       0.000     999.000     999.000
    PRESS91            1.000       0.000     999.000     999.000
    PRESS93            1.000       0.000     999.000     999.000
    PRESS97            1.000       0.000     999.000     999.000
    PRESS00            1.000       0.000     999.000     999.000
    PRESS04            1.000       0.000     999.000     999.000
    PRESS06            1.000       0.000     999.000     999.000
S        |
    PRESS89            0.000       0.000     999.000     999.000
    PRESS91            1.000       0.000     999.000     999.000
    PRESS93            3.000       0.000     999.000     999.000
    PRESS97            7.630       1.758       4.340       0.000
    PRESS00           11.653       2.560       4.552       0.000
    PRESS04           14.896       3.424       4.351       0.000
    PRESS06           14.502       3.310       4.382       0.000
Means ! 均值自由估计,且跨类别变化
    I                 81.655       0.848      96.288       0.000
    S                  0.452       0.114       3.958       0.000
Intercepts
    PRESS89            0.000       0.000     999.000     999.000
    PRESS91            0.000       0.000     999.000     999.000
    PRESS93            0.000       0.000     999.000     999.000
    PRESS97            0.000       0.000     999.000     999.000
    PRESS00            0.000       0.000     999.000     999.000
    PRESS04            0.000       0.000     999.000     999.000
    PRESS06            0.000       0.000     999.000     999.000
Variances! 增长因子方差为 0
    I                  0.000       0.000     999.000     999.000
    S                  0.000       0.000     999.000     999.000
Residual Variances
    PRESS89           72.893       5.116      14.249       0.000
    PRESS91           83.050       6.552      12.675       0.000
    PRESS93           84.123       6.207      13.553       0.000
    PRESS97           96.533       7.867      12.271       0.000
    PRESS00          108.303       8.846      12.243       0.000
    PRESS04          150.587      15.086       9.982       0.000
    PRESS06          150.281      18.378       8.177       0.000
```

```
Latent Class 3
I          |
    PRESS89           1.000      0.000    999.000    999.000
    PRESS91           1.000      0.000    999.000    999.000
    PRESS93           1.000      0.000    999.000    999.000
    PRESS97           1.000      0.000    999.000    999.000
    PRESS00           1.000      0.000    999.000    999.000
    PRESS04           1.000      0.000    999.000    999.000
    PRESS06           1.000      0.000    999.000    999.000
S          |
    PRESS89           0.000      0.000    999.000    999.000
    PRESS91           1.000      0.000    999.000    999.000
    PRESS93           3.000      0.000    999.000    999.000
    PRESS97           7.630      1.758      4.340      0.000
    PRESS00          11.653      2.560      4.552      0.000
    PRESS04          14.896      3.424      4.351      0.000
    PRESS06          14.502      3.310      4.382      0.000
Means ！ 均值自由估计,且跨类别变化
    I                91.065      0.969     93.982      0.000
    S                 0.786      0.194      4.047      0.000
Intercepts
    PRESS89           0.000      0.000    999.000    999.000
    PRESS91           0.000      0.000    999.000    999.000
    PRESS93           0.000      0.000    999.000    999.000
    PRESS97           0.000      0.000    999.000    999.000
    PRESS00           0.000      0.000    999.000    999.000
    PRESS04           0.000      0.000    999.000    999.000
    PRESS06           0.000      0.000    999.000    999.000
Variances
    I                 0.000      0.000    999.000    999.000
    S                 0.000      0.000    999.000    999.000
Residual Variances
    PRESS89          72.893      5.116     14.249      0.000
    PRESS91          83.050      6.552     12.675      0.000
    PRESS93          84.123      6.207     13.553      0.000
    PRESS97          96.533      7.867     12.271      0.000
    PRESS00         108.303      8.846     12.243      0.000
    PRESS04         150.587     15.086      9.982      0.000
    PRESS06         150.281     18.378      8.177      0.000
```

表5-2的拟合指数提示，信息指数随类别数量的增加而逐渐减少，所以很难依据信息指数来选择最佳模型。在此情况下可以采用碎石图来选择（Petras＆Masyn，2010），BIC的碎石图结果呈现在图5-4上。仅从图上来看，在3处存在拐点，考虑选取3个类别。LMR和BLRT的结果

并不一致，在所有模型中，BLRT均提示显著，而LMR在取4个类别时$p=0.492$，提示保留3个类别。但随着类别数目增加到5个时，其结果出现了边缘显著$p=0.030$，而此时的类别概率差异较大，其中最小的为0.025，对应的$n=20$，综合多方信息最后选择保留3个类别的模型。3个类别LCGM对应的增长因子估计均值见图5-5。

表5-2　LCGM和GMM模型拟合信息汇总

Model		k	G^2/LL	AIC	BIC	aBIC	Entropy	LMR	BLRT	类别概率
LCGM										
	1C	13	−17810.415	35646.829	35708.745	35667.460	—	—	—	1
	2C	16	−17403.802	34839.605	34915.808	34864.996	.753	<.001	<.001	.244/.756
	3C	**19**	**−17237.242**	**34512.484**	**34602.976**	**34542.637**	**.738**	**.004**	**<.0001**	**.054/.484/.461**
	4C	22	−17171.30	34386.604	34491.384	34421.517	.704	.492	<.0001	.571/.260/.045/.124
	5C	25	−17129.434	34308.869	34427.937	34348.544	.729	.030	<.0001	.025/.259/.060/.128/.527
	6C	28	−17114.184	34284.368	34417.724	34328.803	.698	.624	<.0001	.476/.047/.204/.212/.015/.046
GMM										
	1C	16	−17187.121	34406.242	34482.446	34431.634	—	—	—	1
	2C	19	−17118.895	34275.790	34366.282	34305.943	.947	<.001	<.001	.954/.046
	3C	**22**	**−17096.568**	**34237.136**	**34341.916**	**34341.916**	**.871**	**.0206**	**<.001**	**.039/.047/.914**
	4C	25	−17092.110	34234.220	34353.288	34273.895	.849	.506	.666	.047/.039/.898/.015
GMM-free										
	2C	**26**	**−17090.677**	**34233.354**	**34357.185**	**34274.615**	**.677**	**.088**	**<.000**	**.096/.904**
	3C	36	−17061.087	34194.174	34365.632	34251.306	.649	.674	.041	.875/.034/.091
GMM-协变量										
	2C	20	−17117.755	34275.510	34370.765	34307.250	.948	<.001	<.001	.9534/.046
	3C	24	−17095.572	34239.144	34353.450	34277.232	.870	.0485	<.000	.042/.913/.045

注：拟合最优的模型被加粗。

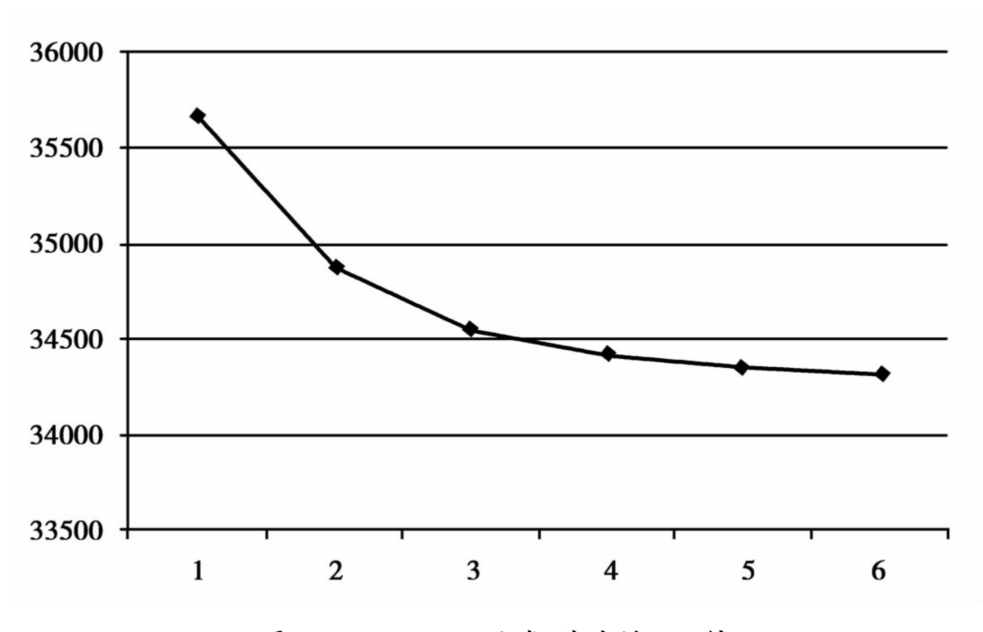


图5-4　LCGM不同类别时的BIC值

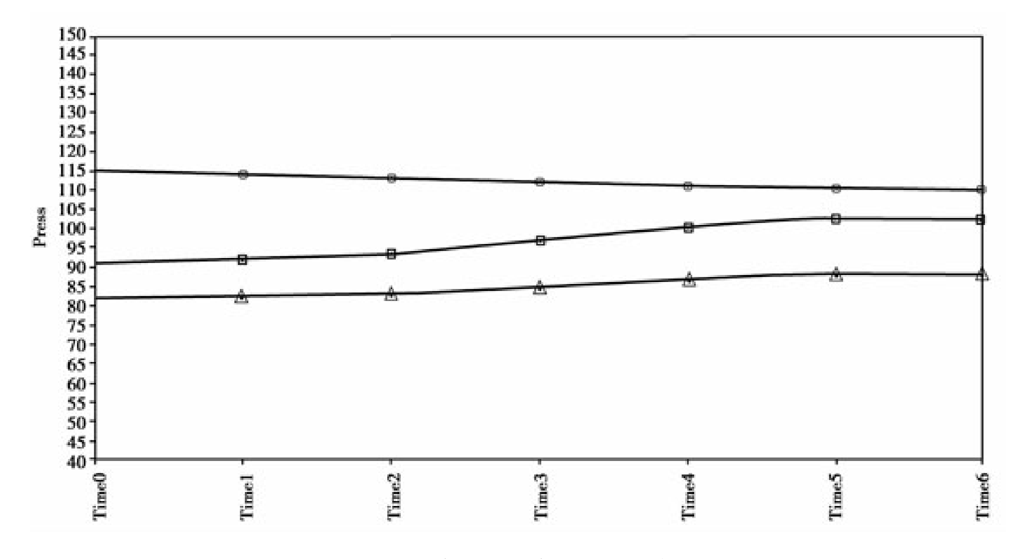


图5-5 LCGM（k ＝3）类别增长轨迹图

如图5-5所示，3个类别组内的个体享有完全相同增长轨迹，并呈现出不同的增长模式。类别组1（图中最上方一组）47人（占全体的5.4％），截距和斜率的均值分别为114.575（p ＜0.01）和－0.308（p ＝0.268），该组的血压起始值最高且呈逐渐下降趋势，将其命名为“高压缓慢下降组”。类别组2（图中中间的一组）419人，截距和斜率的均值分别为81.655（p ＜0.01）和0.452（p ＜0.01），占48.4％，该群体初始血压正常随年龄增长而增加，命名为“缓慢增长组”；类别组3（图中最下方的一组）399人，占46.1％，截距和斜率的均值分别为81.655（p ＜0.01）和0.452（p ＜0.01），命名为“正常组”。

图5-6呈现了3个类别组各自截距因子和斜率因子均值的分布情况。从图中不难发现3个类别组之间存在明显的差异（群体异质性的体现）。

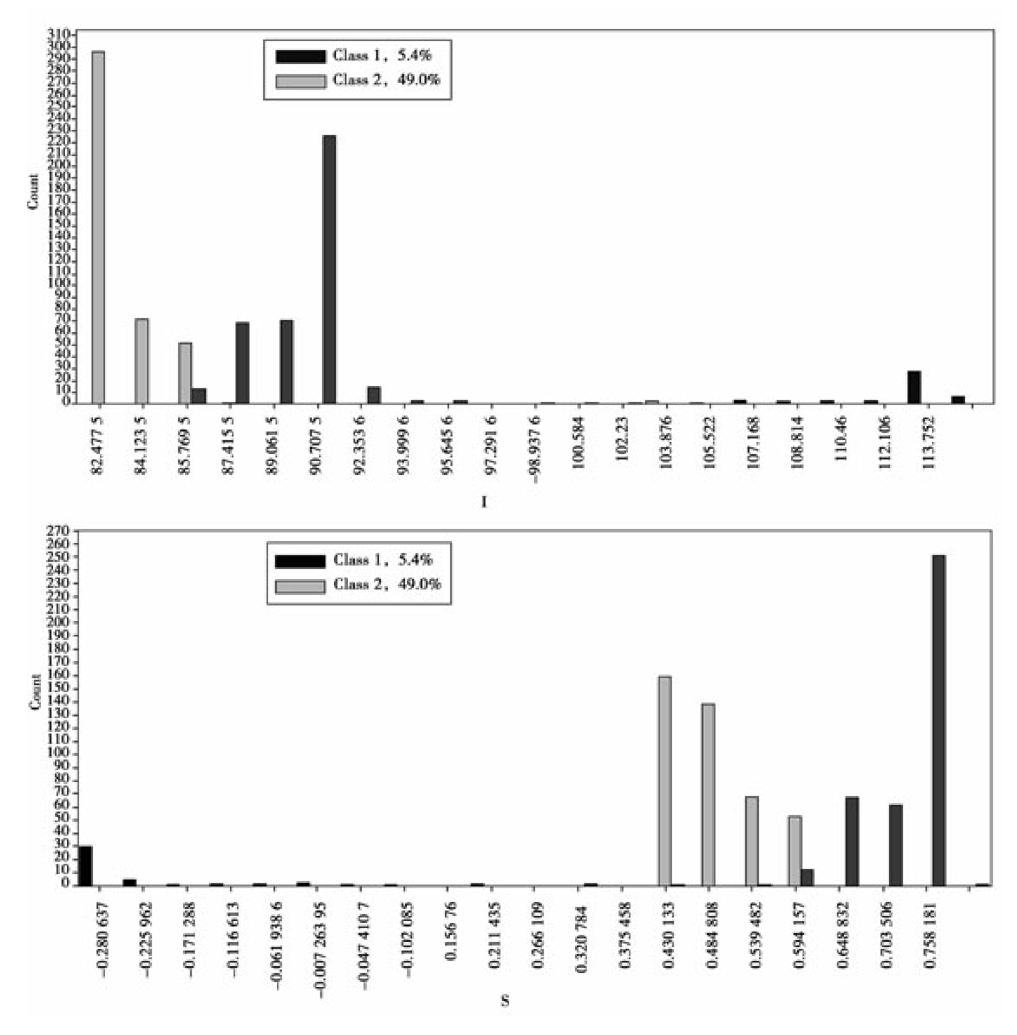


图5-6　潜变量均值分布图（上图是截距因子，下图是斜率因子）

5.5　GMM分析示例

LCGM作为GMM的特例，限定增长因子方差和协方差为0，即假设类别内所有个体具有完全相同的增长轨迹。这一假设比较苛刻，在本例中即使LCGM考虑到群体异质性，增长模型的拟合并没有优化（$\Delta BIC = aBIC_{LCGM} - aBIC_{LGCM} = -111$），下面拟合释放了这一限定的GMM，对应的M*plus* 语句列在框5-2中。

框5-2　GMM语句示例

```
Title: an example of Growth mixture modeling
Data: File is chnspress_wide.dat ;
Variable: Names are id gender bmi89 bmi91 bmi93 bmi97 bmi00 bmi04 bmi06
          press89 press91 press93 press97 press00 press04 press06;
     Missing are all (-9999) ;
USEVARIABLES=press89 press91 press93 press97 press00 press04 press06;
          Classes=c(3) ;
```

```
Analysis: Type=Mixture ;
          Starts=200 10;
          processor=2;
Model:
        %Overall%
          i s | press89@0 press91@1 press93@3 press97 press00 press04 press06;
Output: tech11 tech14;
Plot: Type is Plot3;
     Series=press89 press91 press93 press97 press00 press04 press06(*);
Savedata: SAVE=CPROB;
          FILE IS pressgmm3.txt;
MODEL RESULTS
                                                            Two-Tailed
                         Estimate       S.E.    Est./S.E.     P-Value
Latent Class 1
I          |
     PRESS89                1.000      0.000      999.000     999.000
     PRESS91                1.000      0.000      999.000     999.000
     PRESS93                1.000      0.000      999.000     999.000
     PRESS97                1.000      0.000      999.000     999.000
     PRESS00                1.000      0.000      999.000     999.000
     PRESS04                1.000      0.000      999.000     999.000
     PRESS06                1.000      0.000      999.000     999.000
S          |
     PRESS89                0.000      0.000      999.000     999.000
     PRESS91                1.000      0.000      999.000     999.000
     PRESS93                3.000      0.000      999.000     999.000
     PRESS97                7.309      1.578        4.633       0.000
     PRESS00               10.938      2.298        4.759       0.000
     PRESS04               14.793      3.537        4.182       0.000
     PRESS06               14.346      3.272        4.385       0.000
S          WITH ! 协方差自由估计,但限定跨类别等同。
     I                      0.461      0.370        1.246       0.213
Means ! 均值自由估计,跨组变化
     I                     84.857      2.139       39.675       0.000
     S                      2.385      0.617        3.867       0.000
Intercepts
     PRESS89                0.000      0.000      999.000     999.000
     PRESS91                0.000      0.000      999.000     999.000
     PRESS93                0.000      0.000      999.000     999.000
     PRESS97                0.000      0.000      999.000     999.000
     PRESS00                0.000      0.000      999.000     999.000
     PRESS04                0.000      0.000      999.000     999.000
```

PRESS06	0.000	0.000	999.000	999.000

Variances！方差自由估计,且跨组等同				
I	36.011	4.396	8.193	0.000
S	0.051	0.048	1.052	0.293
Residual Variances				
PRESS89	59.106	4.439	13.314	0.000
PRESS91	71.550	5.446	13.138	0.000
PRESS93	73.587	6.219	11.833	0.000
PRESS97	87.361	7.332	11.914	0.000
PRESS00	88.526	8.166	10.841	0.000
PRESS04	102.046	14.383	7.095	0.000
PRESS06	108.130	13.419	8.058	0.000
Latent Class 2！与类别 1 内容重复的参数未保留				
Means				
I	86.461	0.388	222.916	0.000
S	0.535	0.133	4.014	0.000
Latent Class 3				
Means				
I	115.845	2.878	40.253	0.000
S	-0.831	0.361	-2.302	0.021

由于存在多个类别组，各类别组内的方差和协方差可以设定相等即跨类别组等同（类似测量等值），也可以设定自由估计，限于篇幅下面只报告设定跨组等同的GMM结果，变化的设定读者可以自行尝试。

依次抽取4个类别GMM的拟合结果整理在表5-2中。LMR和BLRT的结论与信息指数的结论并不一致。LMR和BLRT的结果一致提示，选取3个类别的结果优于4个类别，但信息指数则随类别数量的增加而减少。取3个类别时的分类精确性指标*Entropy* ＝0.871优于4个类别的结果0.849。同时，类别概率也较合理，所以综合来看选取3个类别的模型是合理的。

与LCGM相比，GMM各项拟合结果均有显著优化，其中ΔAIC＝275.348，Δ*a* BIC＝200.721，Entropy值也从0.738增加至0.871，分类概率也发生了变化。

3个类别GMM增长趋势图见图5-7，对应的截距和斜率因子估计值呈现在框5-2里。粗线表示模型估计均值，细线表示模型估计的个体值。对比图5-5和图5-7不难发现，两个模型估计的增长模式存在相似之处，但差异更为明显。具体来说，类别组1（图5-7中左下至右上方向一组）起始水平较低，逐步升高，命名为“患病组”，共34人，占全体的3.93％，截距和斜率因子的均值分别为84.857（$p<0.01$）和2.385（$p<0.01$）。

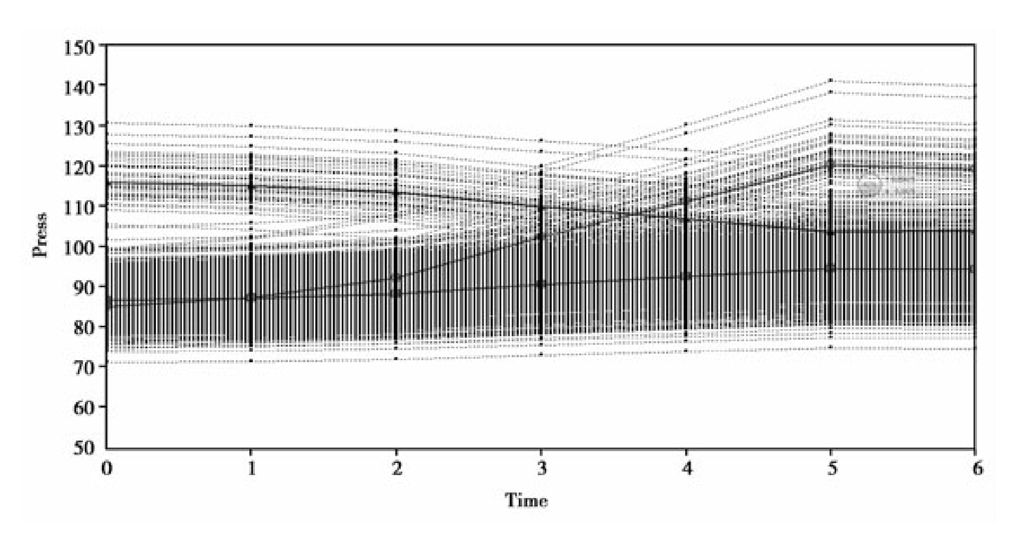


图5-7　3个类别GMM增长趋势图（估计均值和估计个体值）

类别组2（图中左上至右下方向一组），开始水平较高，随时间变化逐渐下降，命名为“康复组”，共40人，占群体的4.62％。该组截距和斜率因子的均值分别为115.844（$p<0.01$）和－0.831（$p=0.021$）。

类别组3（图中最下方一组）始终保持平稳，命名为“正常组”，占全体人数的绝大部分（91.45％），截距和斜率因子的均值分别为84.857（$p<0.01$）和2.385（$p<0.01$）。

截距增长因子和斜率增长因子间的相关（被限定跨类别等值）为0.461（p ＝0.213）[4]，说明初始水平与增长率之间关系不大。截距因子的方差为36.011（p ＜0.01），说明各类别组内的个体间血压初始值存在差异。而斜率因子的方差为0.051（p ＝0.293），说明各类别组内个体间血压增长率差异不显著。

图5-8呈现了3个类别组各自截距因子和斜率因子均值的分布情况。从图中不难发现3个类别组之间存在明显的差异（群体异质性的体现），当然图示的结果与软件输出结果是一致的。另外，对比图5-6不难发现GMM结果与LCGM结果还是存在很大的差异。

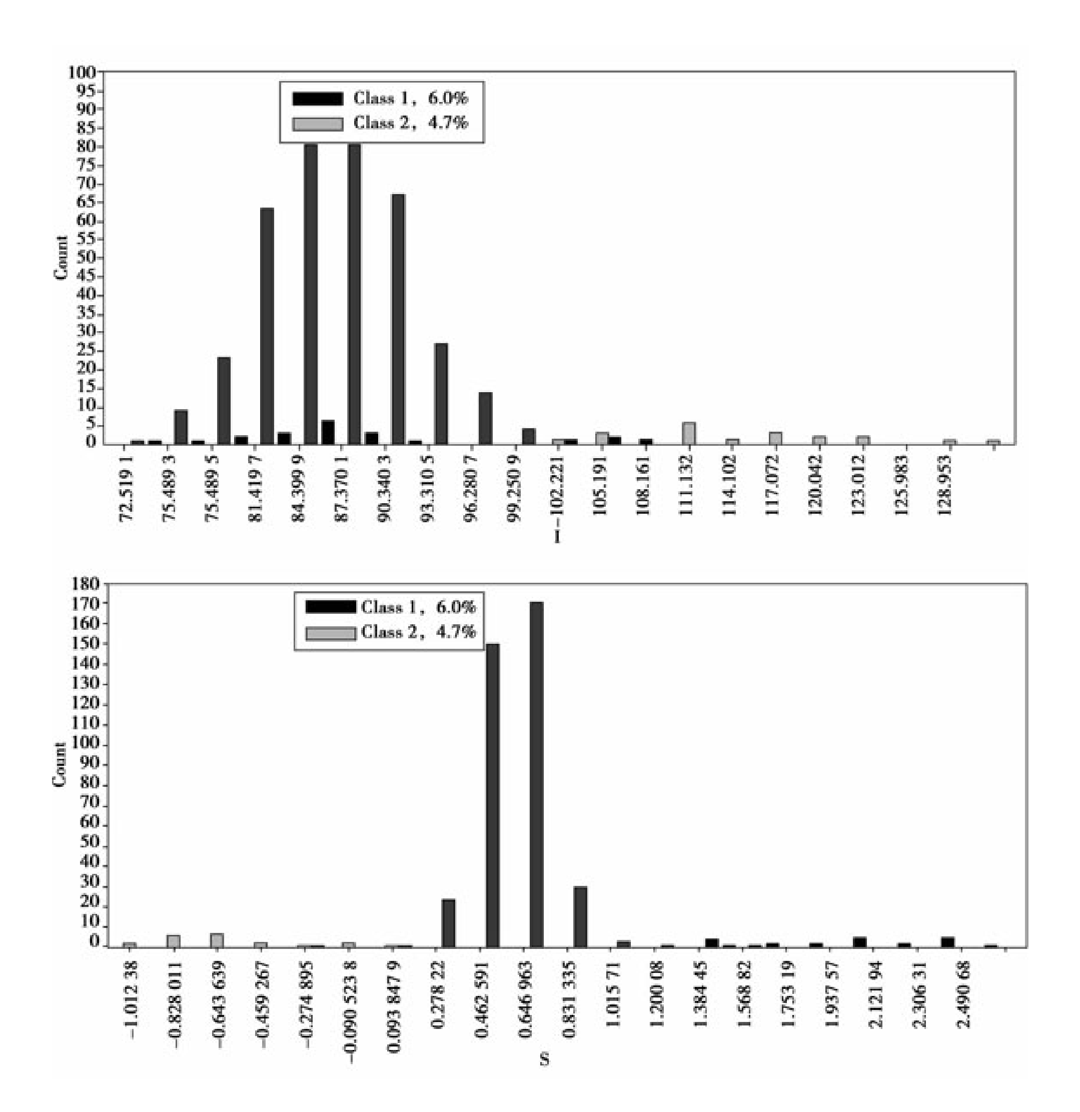


图5-8　GMM潜变量均值分布图（上图是截距因子，下图是斜率因子）

5.6　GMM模型变式

5.6.1　跨类别自由估计的GMM

上述GMM限定了增长因子的方差、协方差及指标测量误差等参数自由估计且限定跨类别等同，但在实际研究中这些设定可能并不合适，所以在实际研究中根据理论预期可以释放某些参数，允许跨类别变化。下面以GMM取2个类别时的模型为例，选择释放增长因子方差和协方差等值限制，允许其跨类别自由估计。M*plus* 语句和部分输出结果见框5-3。

框5-3　跨类别变化的GMM语句

```
Title: an example of GMM with free factor variance & covariance
Data: File is chnspress_wide.dat ;
Variable: Names are id gender bmi89 bmi91 bmi93 bmi97 bmi00 bmi04 bmi06
          press89 press91 press93 press97 press00 press04 press06;
     Missing are all (-9999) ;
     USEVARIABLES= press89 press91 press93 press97 press00 press04 press06;
     Classes=c(2) ;
Analysis: Type=Mixture ;
           Starts=200 10;
          processor=4;
Model:
        %Overall%
     i s | press89@0 press91@1 press93@3 press97 press00 press04 press06;
     %c#1%! 通过在每个类别特定组内设置,解除等值限定。
        i s | press89@0 press91@1 press93@3 press97   press00 press04 press06;
        i with s; ! 协方差自由估计
        i s; ! 因子方差自由估计
     %c#2%
        i s | press89@0 press91@1 press93@3 press97   press00 press04 press06;
        i with s;
        i s;
Output: tech11 tech14;
        Plot: Type is Plot3;
        Series=press89 press91 press93 press97 press00 press04 press06(*);
Savedata: SAVE=CPROB;
           FILE IS pressgmm2.txt;
```

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Latent Class 1				
I |				
PRESS89	1.000	0.000	999.000	999.000
PRESS91	1.000	0.000	999.000	999.000
PRESS93	1.000	0.000	999.000	999.000
PRESS97	1.000	0.000	999.000	999.000
PRESS00	1.000	0.000	999.000	999.000
PRESS04	1.000	0.000	999.000	999.000
PRESS06	1.000	0.000	999.000	999.000
S |				
PRESS89	0.000	0.000	999.000	999.000
PRESS91	1.000	0.000	999.000	999.000
PRESS93	3.000	0.000	999.000	999.000
PRESS97	5.158	2.403	2.147	0.032
PRESS00	5.475	2.257	2.426	0.015
PRESS04	2.990	1.883	1.588	0.112
PRESS06	3.131	0.617	5.075	0.000

```
    PRESS06            3.131      0.617      5.075      0.000
S         WITH
    I                -30.800     22.608     -1.362      0.173
Means
    I                 99.379      2.493     39.865      0.000
    S                  1.054      0.523      2.014      0.044
Intercepts
    PRESS89            0.000      0.000    999.000    999.000
    PRESS91            0.000      0.000    999.000    999.000
    PRESS93            0.000      0.000    999.000    999.000
    PRESS97            0.000      0.000    999.000    999.000
    PRESS00            0.000      0.000    999.000    999.000
    PRESS04            0.000      0.000    999.000    999.000
    PRESS06            0.000      0.000    999.000    999.000
Variances
    I                195.398     57.957      3.371      0.001
    S                 10.203      9.855      1.035      0.301
Residual Variances
    PRESS89           53.284      5.100     10.448      0.000
    PRESS91           72.560      5.813     12.483      0.000
    PRESS93           70.579      7.387      9.554      0.000
    PRESS97           75.010     10.467      7.166      0.000
    PRESS00           78.195      8.801      8.885      0.000
    PRESS04          106.339     15.506      6.858      0.000
    PRESS06          106.205     13.710      7.747      0.000
Latent Class 2
! 为了节省空间,和 class1 重复的内容已删去
S         |
    PRESS89            0.000      0.000    999.000    999.000
    PRESS91            1.000      0.000    999.000    999.000
    PRESS93            3.000      0.000    999.000    999.000
    PRESS97           10.281      4.564      2.252      0.024
    PRESS00           16.047      6.944      2.311      0.021
    PRESS04           24.268     10.474      2.317      0.021
    PRESS06           24.526     10.181      2.409      0.016
```

```
S         WITH
    I                  0.276      0.247      1.119      0.263
  Means
    I                 85.803      0.419    204.716      0.000
    S                  0.398      0.178      2.240      0.025
  Variances
    I                 28.092      3.941      7.129      0.000
    S                  0.090      0.082      1.087      0.277
```

通过在每个类别特定组“％c＃％”下面设置估计的语句，以解除相应的等值限定。具体在本例中，“％c＃％”下面分别增加了增长模

型，增长因子的方差和协方差，在框5-3的下半部分呈现了跨类别自由估计的结果。

从表5-2的拟合指数来看，LMR和BLRT的结论与信息指数的结论并不一致，但LMR和BLRT的结果倾向于支持2个类别的模型。信息指数的结论也是不一致的，AIC和aBIC支持3个类别模型，但BIC支持2个类别模型。2类别的分类精确性也优于3个类别的。综合所有信息后，最终选择2类别模型。

与等值限定的模型相比，允许增长因子的方差和协方差跨类别变化的模型拟合更好，其中2个类别模型的ΔAIC＝42.436，Δa BIC＝31.328，但该模型的分类精确性并不高，Entropy只有0.677，远低于限定等值模型的0.947。

因子方差协方差自由估计的2类别和3类别GMM估计因子均值和个体值如图5-9。从图中不难发现，3类别模型中的第3个类别在后4个时间点稳步上升。

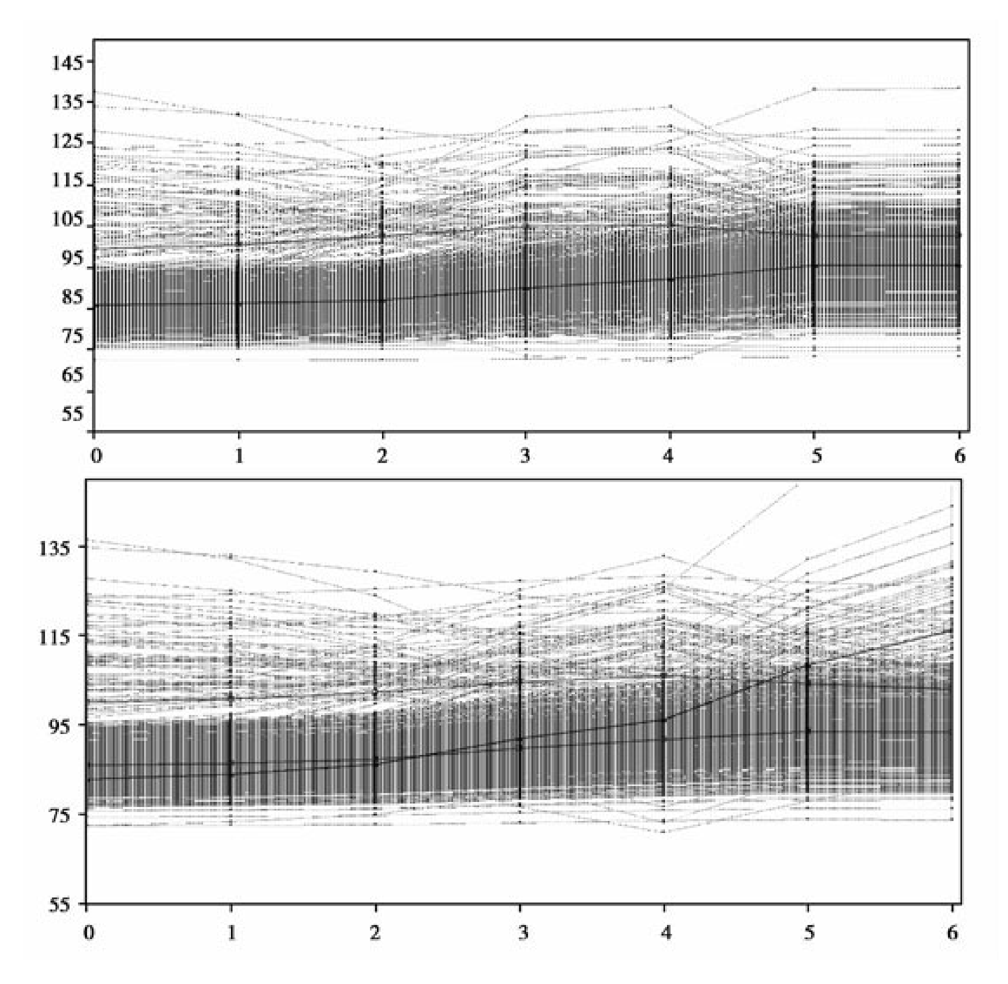


图5-9　因子方差协方差自由估计的GMM（2个和3个类别）：估计的因子均值和个体值

5.6.2　带有协变量的GMM

在实践中可以纳入协变量（如人口学变量）进行混合模型分析，以提高分类精确性（Petras＆Masyn，2010；Nylund et al，2007）。这里使用性别作为预测协变量，其中0代表女性，1代表男性，对应的M*plus* 语句呈现在框5-4中。

从表5-3的拟合结果来看，包含协变量的GMM与未包含协变量的GMM（2类别）之间存在一定的差异但结果出现了不一致，其中ΔAIC＝－2.008，ΔaBIC＝64.684。信息指数支持保留3个类别，而LMR支持选择2个类别，此时的分类精确性也较高。为了与无条件模型进行比较，这里选择报告3类别的结果，模型估计的拟合信息见表5-2。

从模型分类概率来看，加入协变量后，3个类别的比例有所变化，说明协变量的加入确实会影响模型的结果，而且对类别数目的保留也有所影响。具体到本例来说，加入协变量后，取3个类别模型的LMR的p值由0.0206增加到0.0485，使得LMR更倾向于选择2个类别模型。对各类别增长因子的均值和方差也有一定程度的影响（见框5-4）。

框5-4　带有协变量的GMM

```
Title: an example of Growth mixture modeling
Data: File is chnspress_wide.dat ;
Variable: Names are id gender bmi89 bmi91 bmi93 bmi97 bmi00 bmi04 bmi06
          press89 press91 press93 press97 press00 press04 press06;
   Missing are all (-9999) ;
USEVARIABLES= press89 press91 press93 press97 press00 press04 press06;
      Classes=c(3) ;
Analysis:
        Type=Mixture ;
        Starts=200 10;
        processor=4;
Model:
        %Overall%
         i s | press89@0 press91@1 press93@3 press97 press00 press04 press06;
         c on gender;! 增加协变量对类别潜变量 C 的回归效应。
```

```
Output: tech14 tech11;
        Plot: Type is Plot3;
        Series=press89 press91 press93 press97 press00 press04 press06(*);
Savedata: SAVE=CPROB;
         FILE IS pressgmm4.txt;
```

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Latent Class 1				
I					
PRESS89	1.000	0.000	999.000	999.000
PRESS91	1.000	0.000	999.000	999.000
PRESS93	1.000	0.000	999.000	999.000
PRESS97	1.000	0.000	999.000	999.000
PRESS00	1.000	0.000	999.000	999.000
PRESS04	1.000	0.000	999.000	999.000
PRESS06	1.000	0.000	999.000	999.000
S					
PRESS89	0.000	0.000	999.000	999.000
PRESS91	1.000	0.000	999.000	999.000
PRESS93	3.000	0.000	999.000	999.000
PRESS97	7.367	1.624	4.536	0.000
PRESS00	11.059	2.390	4.628	0.000
PRESS04	15.051	3.700	4.068	0.000
PRESS06	14.573	3.422	4.259	0.000
S WITH				
I	0.425	0.384	1.107	0.268
Means				
I	84.923	2.303	36.872	0.000
S	2.343	0.620	3.780	0.000
Variances				
I	35.821	4.333	8.266	0.000

```
    S                    0.053      0.050      1.062      0.288
Latent Class 2！ 与 class1 相同的参数结果已删去
  Means
    I                   86.467      0.389    222.252      0.000
    S                    0.525      0.135      3.906      0.000
Latent Class 3！ 与 class1 相同的参数结果已删去
  Means
    I                  115.845      2.791     41.502      0.000
    S                   -0.785      0.355     -2.209      0.027
Categorical Latent Variables！ 最后一个类别组为参照组
```

```
C#1          ON
    GENDER               0.430      0.636      0.677      0.498
C#2          ON
    GENDER               0.534      0.431      1.239      0.215
Intercepts
    C#1                  0.060      0.557      0.108      0.914
    C#2                  2.710      0.291      9.326      0.000
LOGISTIC REGRESSION ODDS RATIO RESULTS
Categorical Latent Variables
C#1        ON
    GENDER               1.538
C#2        ON
    GENDER               1.705
ALTERNATIVE PARAMETERIZATIONS FOR THE CATEGORICAL LATENT VARIABLE REGRESSION！
采用不同的参照组所得多项式回归的结果。
Parameterization using Reference Class 1
C#2        ON
    GENDER               0.103      0.474      0.218      0.827
C#3        ON
    GENDER              -0.430      0.636     -0.677      0.498
Intercepts
    C#2                  2.650      0.484      5.474      0.000
    C#3                 -0.060      0.557     -0.108      0.914
Parameterization using Reference Class 2
C#1        ON
    GENDER              -0.103      0.474     -0.218      0.827
C#3        ON
    GENDER              -0.534      0.431     -1.239      0.215
Intercepts
    C#1                 -2.650      0.484     -5.474      0.000
    C#3                 -2.710      0.291     -9.326      0.000
```

框5-4还给出了性别预测类别归属的回归系数，并且给出了不同的参数化解。

5.7　本章小结

本章主要介绍了LGCM与LCA相结合的模型GMM以及各种变式。GMM作为新近发展的潜类别增长模型逐渐引起国际学术界的关注，应用研究和方法学研究越来越多。

需要指出的是，尽管GMM近年来引起多方关注，但对GMM的认识并没有像其他潜变量建模方法那样深入全面，GMM作为不断发展的新领域，人们对它的认识也在不断加深，当然也存在一些争论（e．g．，Bauer＆Curran，2003，Hipp＆Bauer，2006）。在将来的研究中，方法学者应该着重关注如下方面的问题：

①全面探讨缺失率和缺失类型对GMM结果（如，潜类别个数保留、参数估计）可靠性的影响。尽管有个别方法学研究涉及了缺失值问题（Kim，2012），然而尚未有研究全面系统探讨不同缺失率及缺失类型对GMM收敛、潜类别数量保留等的影响。

②各种潜类别个数保留指数在不同建模条件下的表现。与结构方程模型的这种拟合指数表现的研究相比，当前对各种GMM类别保留指数表现的研究相当稀少，而且结果也不一致（*P* eugh＆Fan，2012；Nylund，Asparouhov＆Muthén，2007；Tofighi＆Enders，2008；Yang，2006）。

③对GMM结果精确性影响的其他方面。例如样本量（Kim，2012）和模型中是否包含协变量（Li＆Hser，2011）。

总之，GMM作为潜类别模型和增长曲线模型更一般的数据建模方法在分析追踪数据发展趋势的同时兼顾群体异质性，在短短的十几年里

得到研究者的广泛关注。随着统计分析软件的普及，GMM将会被更多的研究者应用。

---

(1) 数据来源：中国健康与营养调查（China Health and Nutrition Survey，CHNS）部分样本。

(2) 通常也称作一般增长混合模型（General Growth Mixture Modeling，GGMM），本章主要涉及连续结局变量。

(3) 这里设置的增长模型为自由估计时间参数的模型，即设定前三次的时间分，其他时间点的负荷由数据估计（见第4章）。

(4) 对应的相关系数为0.341，p ＝0.373

第二部分　多水平模型

6 多水平回归模型

6.1　当OLS遇到嵌套数据

回归是最基础也是应用最广泛的统计分析技术。通过回归模型，可以分析自变量解释因变量的影响是否具有统计学上的显著性，以及作用幅度的大小。以简单一元回归为例：

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{6.1}$$

其中，Y_i 为因变量或被解释变量，X_i 为自变量或解释变量，β_0 为截距，β_1 为斜率，ε_i 为误差项。

推广到有k 个自变量的情况，即为一般的多元回归。可用如下公式表示：

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \varepsilon_i \tag{6.2}$$

其中，Y_i 为因变量或被解释变量，X_{i1}，…，X_{ik} 为自变量或解释变量，β_0 为截距，β_{i1}，…，β_{ik} 为斜率或偏回归系数，ε_i 为误差项。

然而要注意的是，线性回归模型只有在满足一系列的假定的条件下，才能得到无偏有效的估计。一般来讲，需要满足如下假定条件：

- 线性假定，某些非线性函数通过转换变为线性函数同样适用。
- 正交假定：①零条件均值，即$E(\mu \mid X_{i1}, X_{i2}, \cdots, X_{ik})=0$；②误差项$\varepsilon_i$ 与X_i 不相关，即Cov（X_i，ε_j）＝0。另可推导，$\mathrm{Cov}(\hat{Y}_i, \varepsilon_j) = 0$。

- 独立同分布假定（i．i．d．）①任意两个误差项之间协方差为0，即Cov（ε_i，ε_j）＝0，$i \neq j$；②同方差性，对于所有的X值，ε的方差σ^2相同，即Var（ε_i）＝Var（ε_j）＝σ^2。
- 正态分布假定，误差项ε是一个服从正态分布的随机变量，且相互独立，即ε～N（0，σ^2）。[(1)]

若满足正交假定和独立同分布假定，最小二乘法得到的回归参数估计是所有估计中最佳线性无偏估计（best linear unbiased estimator）。此即所谓“高斯-马尔科夫定理”。[(2)]

现实当中，我们需要面对不同的研究层次，研究对象嵌套在不同层次的单位之中。比如我们研究学生成绩，学生来自不同的班级，这些班级又嵌套在不同的学校之中。研究企业来自不同的产业区，产业区又位于不同的城市。在进行重复测量的设计时，不同试点的测量嵌套于同一个个体之内（历时数据、增长模型），参数的估计嵌套于不同的研究（meta-analysis）。嵌套（nested）、聚类（cluster-correlated）、分层（hierarchical）、多水平（multilevel）等概念都是用来刻画这样一种数据结构特点。图6-1即属于这种情况，不同的学生嵌套于不同的学校，不同学生又分别测量了k次，历次测量可以看作嵌套于学生个体。

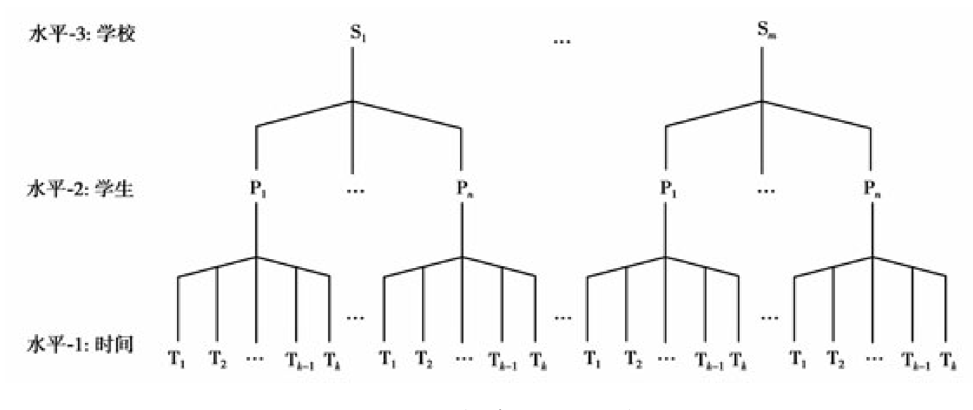


图6-1　嵌套数据结构

实际上，在抽样调查时，由于可以大幅度节省成本，多阶段聚类抽样（multi-stage cluster sampling）十分常用，或者进行历时研究，在不同时点对不同的个体进行多次测量，都自然会形成嵌套数据结构。也就是说，嵌套数据结构在现实当中是非常普遍的。

使用一般的回归模型处理嵌套数据存在诸多局限。不同的研究对象由于嵌套于不同的时空单位，彼此之间更具相关性。比如来自同一家庭的兄弟姊妹，其相似性肯定比来自不同家庭的个体之间要高。这破坏了使用OLS回归残差相互独立的假定。在不存在内生问题的情况下，虽然回归系数估计仍然无偏，但标准误的估计偏小，尤其对于组群层面变量，导致统计效力损失，因而所犯第一类错误的概率往往偏大。许多显著结果可能是值得怀疑的。而且，如果忽略嵌套结构，显然也未充分利用数据中包含的信息。

因果推论时，如果分析单位和推论单位属于不同的层次，用某一层数据分析，而在另外的层上得出结论，在方法论上，就会造成“生态谬误”（ecological fallacy）或“还原谬误”（atomistic fallacy）的问题（Hox，2010）。[3]

统计上，对于嵌套于不同组群内部的个体之间的相似性，可以使用所谓“组内相关系数”（intra-class correlation coefficient）来精确衡量。ICC（也以 ρ 表示）定义为组间方差占总方差的比例：

$$\mathrm{ICC} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_w^2} \tag{6.3}$$

其中，σ_b^2 为组间方差（between-group variance），σ_w^2 为组内方差（within-group variance），$\sigma_b^2+\sigma_w^2$ 为总方差（total variance）。

组内同质即组间异质。因此，ICC既然反映组间变异，其实另一面就是组内个体相关。其取值理论上为0到1。如果出现负值，可能是由于模型上设定错误或组间方差接近于零所致。ICC趋近于0，表示没有组群效应，组内个体趋于相互独立。

如果ICC很小，此时模型可使用一般回归模型进行估计，无需进行多水平模型分析。否则，传统回归模型观测值独立假设被破坏，需进行多水平分析。不同研究领域ICC标准差异很大。Cohen（1977）指出，ICC小于0.059时，属于小的组内相关；介于0.059和0.138之间时，属于中等相关；高于0.138，属于高度组内相关。中等程度组内相关就不能忽略组内相似性的存在，因此ICC大于0.059时，就需要使用多水平模型进行分析。决定是否使用多水平分析，也要考虑设计效应（design effect）。

$$deff = 1 + (\text{组群平均规模} - 1) \times \rho \tag{6.4}$$

当设计效应大于2时，考虑进行多水平分析（Satorra＆Muthén，1995）。

目前常见的多水平分析软件主要包括：HLM、M*plus* 、MLwin、SAS（PROC MIXED）、Stata（xtmixd，gllamm）、SPSS（MIXED）、R（multilevel，nlme）等。对于嵌套数据的处理，在这些软件中，M*plus* 的优势在于可以允许因变量是连续变量、二分变量、有序或无序类别变量、截取变量、计数变量等多种类型，而且还可以与路径分析、结构方程模型、因子分析、潜类分析等相结合，将这些模型置于多水平的框架之下。

需要注意的是，不同的软件所使用的多水平数据结构组织上有区别。HLM要求分别输入个体水平和组群水平的数据集，而M*plus* 、SAS、SPSS、R、MLwin等软件则要求输入混合数据集。

我们看一个实例。PISA（*P* rogramme for International Student Assessment）是经合组织（OECD）1997年开始的一项国际调查，每三年一次，在世界范围内通过对15岁青少年阅读、数学和科学的测验评估教育系统状况。目前有70个国家和经济体参与了PISA项目。更多信息可参阅官方网站。后续分析以美国2009年数据[(4)]为例，涉及的主要变量如下：

- 水平-1文件（INT__STQ09__DEC11．sav）：
  HISEI，父母最高的社会境地地位指数（SEI）；
  FEMALE（ST04Q01），学生性别（1＝男生，0＝女生）；
  PV1MATH，数学成绩似真值（plausible value）；
  PV1READ，阅读成绩似真值；
  PV1SCIE，科学成绩似真值；
- 水平-2文件（INT__SCQ09__Dec11．sav）：
  SECTOR（SC02G01），学校类型（1＝私立，0＝公立）；

SCHSIZE，学生规模；PROPQUAL，有教师证的教师比例。

该数据属于嵌套数据。其中，图6-2、图6-3分别为学生个体层面和学校层面的数据，而图6-4显示的是学生个体层面和学校层面的混合数据。

	schoolid	hisei	pv1math	pv1read	pv1scie	female
7	1	51	386.13	437.78	434.39	1
8	1	34	527.11	493.62	489.13	1
9	1	21	355.83	405.13	423.85	1
10	1	34	351.07	321.52	365.39	0
11	1	23	406.46	412.04	437.75	1
12	1	69	377.64	394.95	447.35	0
13	1	43	568.48	566.39	595.43	1
14	1	51	659.92	666.32	675.81	1
15	1	61	523.3	593.69	563.63	0
16	1	51	540.82	615.48	606.53	1
17	1	43	420.48	393.03	409.12	0
18	1	70	512.78	453.5	544.33	1
19	1	67	589.04	630.33	642.24	0
20	1	34	393.6	320.48	399.52	0
21	1	71	446.49	444.34	513.56	0
22	1	34	608.9	604.84	601.03	1
23	1	34	333.32	355.24	360.44	1
24	1	38	506.24	665.05	630.87	1
25	2	53	420.09	425.58	436.16	0
26	2	30	510.45	533.8	520.09	0
27	2	55	491.67	459.17	489.31	0
28	2	70	440.73	488.99	471.41	0

图6-2　水平-1：学生数据库

一般来讲，原始数据格式不同层面都是分开的，在使用*Mplus*软件进行多水平分析之前，需要进行合并。在Stata、SPSS等具有数据管

	schoolid	propqual	schsize	sector
1	1	1	808	0
2	2	.616	899	0
3	3	1	996	0
4	4	1	1273	0
5	5	9999	929	0
6	6	1	2004	0
7	7	1	3592	0
8	8	1	2640	0
9	9	1	1680	0
10	10	1	149	0
11	11	1	247	0
12	12	1	1574	0
13	13	1	99999	0
14	14	1	1101	0
15	15	1	2144	0

图6-3　水平-2：学校数据库

	schoolid	female	hisei	pv1math	pv1read	pv1scie	sector	schsize	propqual
10	1	0	34	351.07	321.52	365.39	0	808	1
11	1	1	23	406.46	412.04	437.75	0	808	1
12	1	0	69	377.64	394.95	447.35	0	808	1
13	1	1	43	568.48	566.39	595.43	0	808	1
14	1	1	51	659.92	666.32	675.81	0	808	1
15	1	0	61	523.3	593.69	563.63	0	808	1
16	1	1	51	540.82	615.48	606.53	0	808	1
17	1	0	43	420.48	393.03	409.12	0	808	1
18	1	1	70	512.78	453.5	544.33	0	808	1
19	1	0	67	589.04	630.33	642.24	0	808	1
20	1	0	34	393.6	320.48	399.52	0	808	1
21	1	0	71	446.49	444.34	513.56	0	808	1
22	1	1	34	608.9	604.84	601.03	0	808	1
23	1	1	34	333.32	355.24	360.44	0	808	1
24	1	1	38	506.24	665.05	630.87	0	808	1
25	2	0	53	420.09	425.58	436.16	0	899	.616
26	2	0	30	510.45	533.8	520.09	0	899	.616
27	2	0	55	491.67	459.17	489.31	0	899	.616
28	2	0	70	440.73	488.99	471.41	0	899	.616
29	2	0	37	522.83	451.55	430.57	0	899	.616
30	2	1	70	415.8	448.58	482.51	0	899	.616

图6-4　混合数据库

一般来讲，原始数据格式不同层面都是分开的，在使用*Mplus*软件进行多水平分析之前，需要进行合并。在Stata、SPSS等具有数据管

理功能的统计软件中，使用合并（merge）过程可以很容易实现。对于上述数据，由于学生数据中有其所属学校的编号（shoolid），可以此为索引变量，将学校数据合并进来，形成混合数据。

6.2　多水平回归模型的优势

当然，如果不使用多水平回归模型，也有一些替代性的方法。对于嵌套结构的数据，最简单的处理方式可能就是忽略群组层面变量，只用当前层面的变量进行分析。比如对于PISA数据，在研究学生个体的数据中数学成绩决定因素时，只纳入性别、家庭社会经济地位等个体层面变量。然而，这样没有解决由于数据嵌套而带来的违反误差项独立假定的问题，还因为忽略重要的学校属性变量而可能导致出现误差项与自变量相关的内生性（endogenity）问题，造成系数估计有偏。

这种完全忽视组间变异的模型也称作整体或混合回归（total or pooled regression，Kreft＆De Leeuw，1998）。从表6-1可看到，如果只考虑个体层面的变量，省略重要自变量包括水平-2变量，家庭社会经济地位（hisei）对数学成绩回归系数的估计有偏（参照其他模型，可以看到这里是高估），而标准误则被低估，这容易造成拒绝原假设接受研究假设的结果。

表6-1　整体回归、脉络回归与Cronbach回归

	(1) pooled	(2) contextual	(3) Cronbach
hisei	1.814*** (0.0740)	1.133*** (0.0783)	
grouphisei		3.8033*** (0.185)	
cgrouphisei			1.133*** (0.0783)
group_grancl_hisei			4.935*** (0.168)
_cons	394.6*** (3.994)	233.6*** (8.724)	488.1*** (1.135)
N	4842	4842	4842
R^2	0.110	0.182	0.182
R_a^2	0.110	0.181	0.181

括号内为标准误

p <0.10，* p <0.05，** p <0.01，*** p <.001

在回归模型3中加入水平-2的各学校学生家庭社会经济地位均值（grouphises）变量，得到脉络回归（contextual regression）。将水平-2变量直接纳入水平-1属于聚解（disaggregation）方式，显然没有考虑数据嵌套结构，误差项相互独立的假设被破坏，回归系数标准误仍然被低估，而且可能会引入多重共线性。

Cronbach模型同时在模型中纳入经过中心化的个体效果（$X_{ij}-X_{.j}$）和组群团体效果（$X_{.j}-X_{..}$），二者为正交关系。该模型解决了个体层次与聚合变量之间存在的共线性问题，但回归系数标准误被低估，原因在于组内相关以及回归系数标准误计算观测值并非真正的组群数（Cronbach，1976；Kreft＆De Leeuw，1998）。

对于嵌套数据，另一种处理方式是将所有变量汇总到群组层面，从研究学生到研究学校，然后使用一般回归模型，即所谓组间效应模

型（between effect model）或聚合模型（aggregated regression）。在PISA数据中，可将数学成绩得分聚合到水平-2，然后对各学校数学平均成绩和学生平均社会经济地位进行回归分析。

这种模型的变量虽然均处于同一层次，避免了使用一般模型估计嵌套数据存在的问题，但由于分析层次的变换，完全忽略了组内变异，损失了大量水平-1单位信息，而且在群组层面所得结论并不能推论到个体层面，否则会出现“生态谬误”的问题。比如，在国家层面，富裕的国家对社会福利支持更高，但不能由此推论富人更支持社会福利。也就是说富裕在组间和组内有相反的效应。这样的结果通常是遗漏水平-1的重要变量引起的。如果控制了政治保守性，两个层面效应可能会一致。

从表6-2结果来看，在学校层面，学生家庭社会经济地位指数平均值每增加1分，学校学生数学成绩平均分增加5分左右，远远超出学生个体层面家庭社会社会经济地位指数与数学成绩之间的系数估计。实际上，这本来也说的是两个不同层面的问题。

表6-2　聚合模型

	(1)	(2)
grouphisei	5.013*** (0.370)	5.190*** (0.400)
sector		−20.23+ (11.47)
schsize		−0.000661*** (0.000212)
stratio		0.00566* (0.00222)
propqual		−0.00292 (0.00194)
_cons	228.9*** (19.05)	221.5*** (20.40)
N	152	152
R^2	0.551	0.592
R^2_a	0.548	0.578

括号内为标准误

$p<0.10$，* $p<0.05$，** $p<0.01$，*** $p<0.001$

变量grouphisei是通过水平-1变量hisei的聚合（aggregation）而得。所谓聚合即样本构成特征或样本内某类型个体汇总（一般为均值或百分比），不过水平-2师资水平、硬件设施等信息属于整体特征（global characteristics），只能在群组层面收集。在一些文献中，前者称为“情境变量”（contextual variables），后者称为“整体变量”（global或integral variables）（Blakely & Woodward，2000；Enders＆Tofighi，2007）。[5]

加入区别不同组群的虚拟变量（dummy variables）到回归模型中，构成所谓“虚拟变量回归”，每个组群有不同的截距，从而吸收了组群间的变差（也包括相关的误差项）。这样，每一组群都会得到不同的截距，除了误差项外，其他变量与因变量的关系只是在组群内解读。

$$Y_{ij} = \beta X_{ij} + \sum_{j=1}^{n} \alpha_j D_{ij} + \varepsilon_{ij} \quad (6.5)$$

$$D_{ij} = \begin{cases} 1 & \text{观测值来自第 } j \text{ 组} \\ 0 & \text{其他} \end{cases}$$

这种方法简单易行，但如果组群数比较多，模型过于繁琐，结果解读也有困难。而且由于模型主要关注组内变异，如果组内变异不大，或更关注组间变异，这种方式不适用。这种处理方式用虚拟变量拟合高水平的固定效应，本质上属于固定效应模型，可用更简单的方式来处理。

固定效应模型（fixed effects model，或within-group effects）是用来处理聚类数据的常用方法之一。固定效应模型：

$$Y_{ij} = \beta X_{ij} + \alpha_j + \varepsilon_{ij} \quad (6.6)$$

其中，i 代表j 个组群中的某个体；如果是重复测量，也可能是j 个体的第i 次测量。α_j 对于每个组群或重复观测个体取值不同，代表非观测的（实际上观测因素也可以包括在内）、各组群的某种整体特征，如国家历史、制度或重复观测个体不随时间而变化的因素（比如能力）。α_j 可与X_{ij} 相关，即Cov（α，X_{ij}）≠0。[6]由于α_j

难以直接观测，FEM一般通过离均差法（mean deviation）的组内（within）转换，消除了α_j，因为$\alpha_j = \bar{\alpha}_j$：

$$Y_{ij} - \bar{Y}_j = \beta(X_{ij} - \bar{X}_j) + (\varepsilon_{ij} - \bar{\varepsilon}_j) \quad (6.7)$$

另一种within转换是为每个群体或重复观测个体加一个指示用的虚拟变量，即“虚拟变量法”。如果是重复观测，也可以用前后时点数据相减，形成一阶差分（first difference）。

转换之后，不管是解释变量还是误差部分，模型看起来只包括组内变差。经过转换可以使用OLS估计回归系数，不过回归系数无偏，但标准误和p值因模型估计时潜在使用了代表样本内中不同个体虚拟变量的个数而可能产生新的问题，所以最好使用某些软件提供的FEM命令进行估计（Allison，2009）[7]。

组内效应（within-effect）与组间效应（between-effect）之间存在什么关系呢？一般来讲，组内效应与组间效应方向多是一致的。比如学习努力的学生成绩好，而那些拥有更多努力学习的学生的班级，平均成绩也好。但大部分时候，变量之间的关系强度在组内和组间是不同的，甚至有时关系方向是相反的。比如富人不支持社会福利，但拥有更多富人、人均收入高的国家，可能是高福利国家。实际上，整体回归系数可以被表达为组内和组间系数的加权平均值，参见图6-5（Snijders＆Bosker，2012：31）。

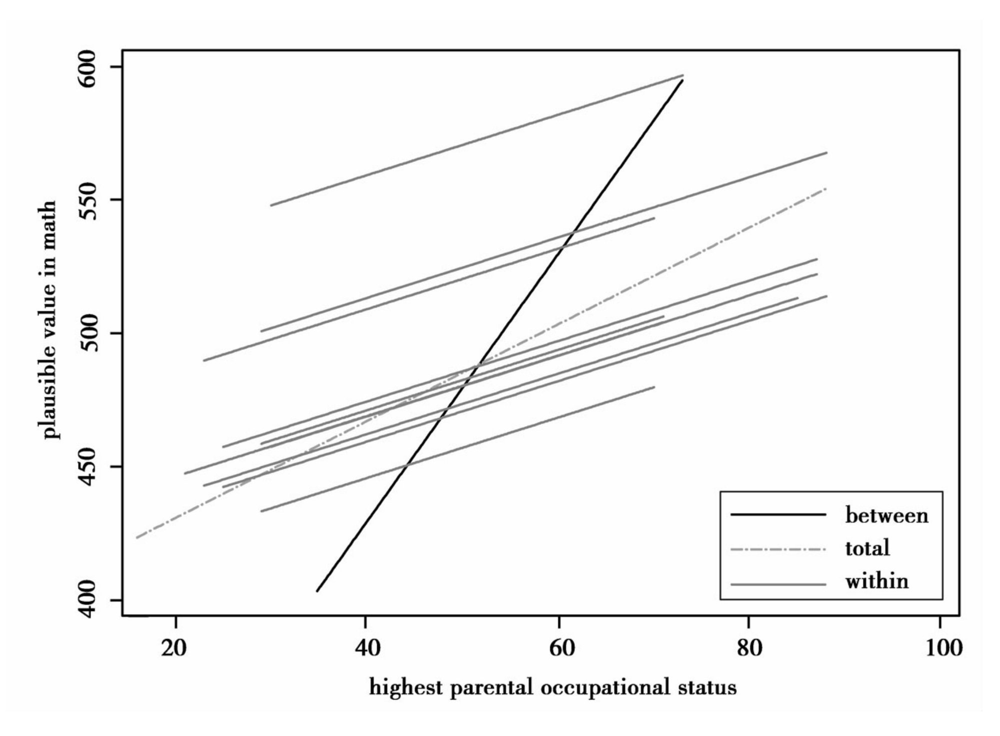


图6-5　组内回归与组间回归模型

从上述模型来看，聚合回归只关注组间变化，组内回归主要关注组内变化。实际上，可以在同一个模型中纳入组间和组内效应。可以将某个变量如SEI，在各组之内取均值，再用组内各SEI值减去组均值。两个变量同时纳入带有随机效应的模型，分别估计组间和组内效应。

表6-3　固定效应模型

	(1) OLS	(2) Dummy	(3) FEM	(4) Within
hisei	1.814*** (0.0740)	1.133*** (0.0734)	1.133*** (0.0734)	1.133*** (0.0722)
2.schoolid		5.424 (19.21)		
3.schoolid		−9.574 (19.63)		
…	…			
cons	394.6*** (3.994)	423.7*** (15.44)	429.7*** (3.931)	0.000 (1.0467)
N	4842	4842	4842	4842
R^2	0.110	0.304		0.048
R^2_a	0.110	0.281		0.048

括号内为标准误

p <0.1，* p <0.05，** p <0.01，*** p <0.001

以PISA数据为例，暂不考虑残差项，针对hisei与pv1math之间的关系，已经估计得到如下回归方程：

整体回归：

$$Y_{ij} = 394.6 + 1.814 \times X_{ij} \tag{6.8}$$

组间回归：

$$\overline{Y}_{ij} = 228.9 + 5.013 \times \overline{X}_{ij} \tag{6.9}$$

组内回归，这里采取OLS估计：

$$(Y_{ij} - \overline{Y}_{ij}) = 1.133 \times (X_{ij} - \overline{X}_j) \quad (6.10)$$

将$\overline{Y}_{ij}$ 移到等号右侧，

$$Y_{ij} = \overline{Y}_{ij} + 429.7 + 1.133 \times (X_{ij} - \overline{X}_j) \quad (6.11)$$

将组间回归方程方程带入上式，

$$\begin{aligned} Y_{ij} &= (228.9 + 5.013 \times \overline{X}_j) + [1.133 \times (X_{ij} - \overline{X}_j)] \\ &= 228.9 + 5.013 \times \overline{X}_j + 1.133 \times X_{ij} - 1.133 \times \overline{X}_j \\ &= 228.9 + 1.133 \times X_{ij} + 3.880 \times \overline{X}_j \end{aligned} \quad (6.12)$$

上式其实即可以看作多水平随机截距模型的固定部分，水平-2截距方程含水平-1变量的聚合变量。这表明，只有将组内模型和组间模型联合考虑也就是采取多水平回归模型，才能揭示这里自变量和因变量的真实关系（Snijders＆Bosker，2012：29）。

多水平回归模型契合嵌套数据结构，与传统的单水平模型相比，多水平回归模型纳入不同层次的因素，同时又保持这些因素在各自的层次，避免了聚合（aggregation）或聚解（disaggregation）带来的问题。可将宏观、微观因素纳入同一个模型，并对跨水平交互作用进行研究。比如，如果假定不同的学校学生有不同的数学成绩得分，那么就可以构造一个截距模型。如果SES对数学成绩的影响在不同的学校也存在不同模式，那就是斜率也发生变化的模型，见图6-6。

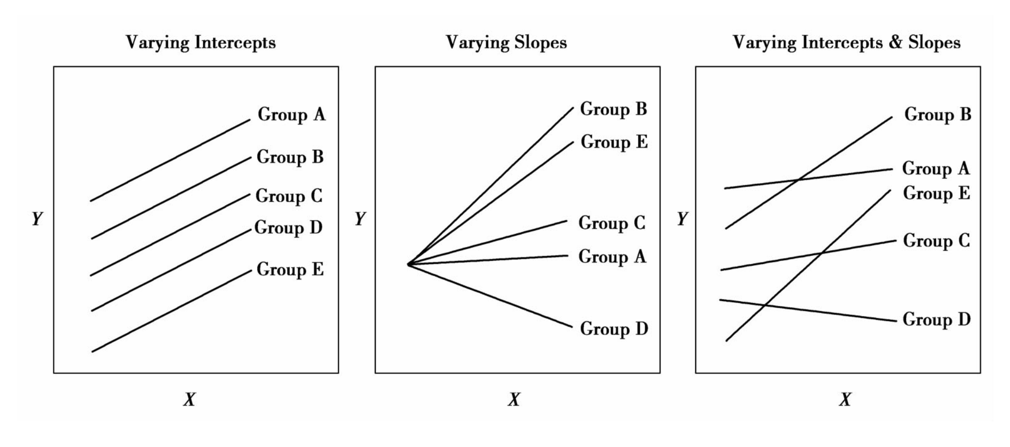


图6-6　变动截距与斜率模型示意图

另外，多水平回归模型据可采取多种估计方法和算法进行估计，具有更好的实用性，有效改进估计的质量，可提供稳健的标准误估计。[(8)]

下一节将就多水平回归模型进行讲解。篇幅所限，内容以两水平和结果变量为连续变量的示例为主。实际上，M*plus* 目前版本已支持三水平建模，同时只需在变量设定部分指定变量的类型，就可以方便地将模型扩展到二分、多分类、定序、计数乃至截取等不同的情况，并且可与潜类分析相结合构成混合模型。这正是M*plus* 强大之处的一个体现。

6.3　一个典型的两水平回归模型

多水平回归模型可视为一般回归模型的拓展。以PISA数据为例，针对所有学校，考虑研究学生SES对数学成绩的影响，在个体层面构建如下模型：

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij} \qquad (6.13)$$

其中，$\varepsilon_{ij} \sim N(0, \sigma^2)$。

每个学校学生SES与数学成绩的关系模式都有所不同，需要建立针对β_{0j}和β_{1j}的模型来刻画。除了一些随机的因素，可用学校的某种特征来预测，比如学校的性质属于私立学校（编码为1）还是公立学校（编码为0）。有人假设私立学校比公立学校平均成绩高（截距大），而就SES对数学成绩的作用而言，在私立学校比在公立学校要小（斜率小）。这种关系可以通过在水平-2建立模型来刻画：

$$\beta_{0j} = \gamma_{00} + \gamma_{01} W_j + u_{0j} \quad (6.14)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} W_j + u_{1j} \quad (6.15)$$

其中，$\begin{bmatrix} \mu_{0j} \\ \mu_{1j} \end{bmatrix} \sim N\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix}\right]$。

方程6.14和方程6.15中，令$E(\beta_{0j}) = \gamma_0$，$Var(\beta_{0j}) = Var(\mu_{0j}) = \tau_{00}$；$E(\beta_{1j}) = \gamma_1$，$Var(\beta_{1j}) = Var(\mu_{1j}) = \tau_{11}$；$Cov(\beta_{0j}, \beta_{1j}) = \tau_{01}$。

各参数均有实际意义：γ_0为学校平均成绩的总体均值；τ_{00}为学校平均成绩的总体方差；γ_1为各校SES与成绩间斜率的总体均值；τ_{11}为各校斜率的总体方差；τ_{01}为斜率和截距间的总体协方差。

方程6.14和6.15中各参数代表的意义为：γ_{00}为公立学校的平均成绩；γ_{01}为私立学校与公立学校在平均成绩上的差异；γ_{10}为公立学校的SES状况对平均成绩作用的平均斜率；γ_{11}为私立学校

与公立学校在SES状况对平均成绩作用斜率上的平均差异；u_{0j} 为在控制W_j 不变时学校j 对平均成绩的独特作用；u_{1j} 为在控制W_j 不变时学校j 在SES对平均成绩作用斜率上的独特作用。

因为β_{0j} 和β_{1j} 不能观测，方程6.14和6.15并不能直接进行估计，将它们代入方程6.13，得复合模型（composite model）或混合模型（mixed model）：

$$Y_{ij} = \underbrace{[\gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}W_j + \gamma_{11}W_jX_{ij}]}_{\text{固定效应}} + \underbrace{[u_{0j} + u_{1j}X_{ij} + \varepsilon_{ij}]}_{\text{随机效应}} \quad (6.16)$$

式6.16随机误差具有复杂的形式，其值在各学校内部不是互相独立的，且各个学校之间也存在差异，因此不能用OLS估计，除非u_{0j} 和u_{1j} 同时为0。[9]

上述模型以图形的方式直观刻画，可参见图6-7，这是一个M*plus* 风格的多水平回归模型图示。图形直观地显示出，在多水平回归模型中，高水平的预测变量对低水平结果变量的影响，是通过间接作用于低水平方程的截距和斜率的影响实现的。

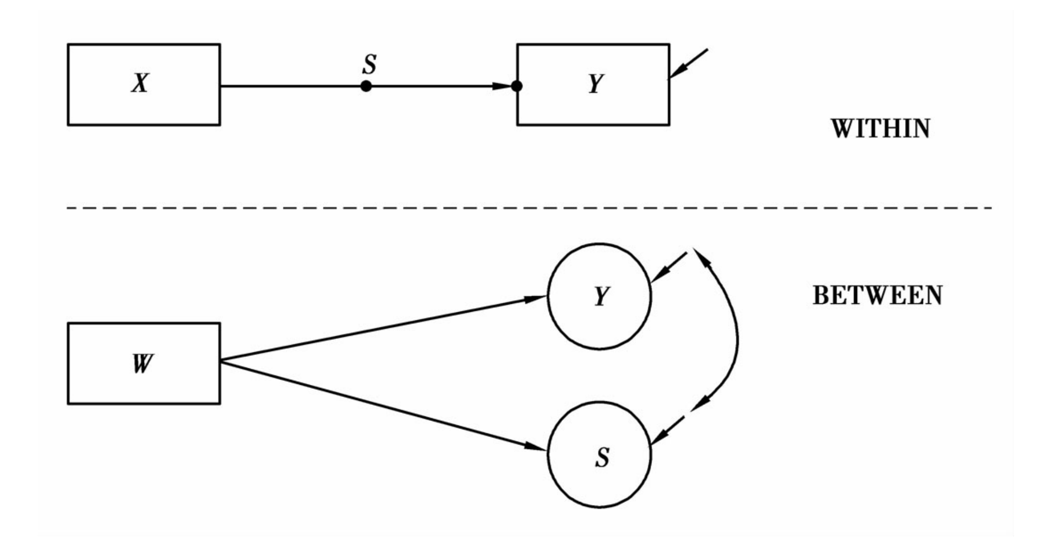


图6-7　两水平模型

图中S代表斜率即β_{1j}，WITHIN部分Y代表水平-1的结果变量，BETWEEN部分Y代表水平-1截距即β_{0j}。

上述方程，对应于所谓“以截距和斜率作为结果的多水平回归模型”（参见本章第7节 ）。模型求解后，采用贝叶斯估计的方式，取得截距和斜率的估计，并将各学校对应的模型回归线制图，参见图6-8。可更清晰地看到，在一些学校之间，学生家庭社会经济地位与数学成绩的关系的确存在一定程度的模式差异。不过图形反映出，就本例而言，不同类型学校之间，数学成绩的截距差异较为明显，相对来说，斜率的变异性要小。

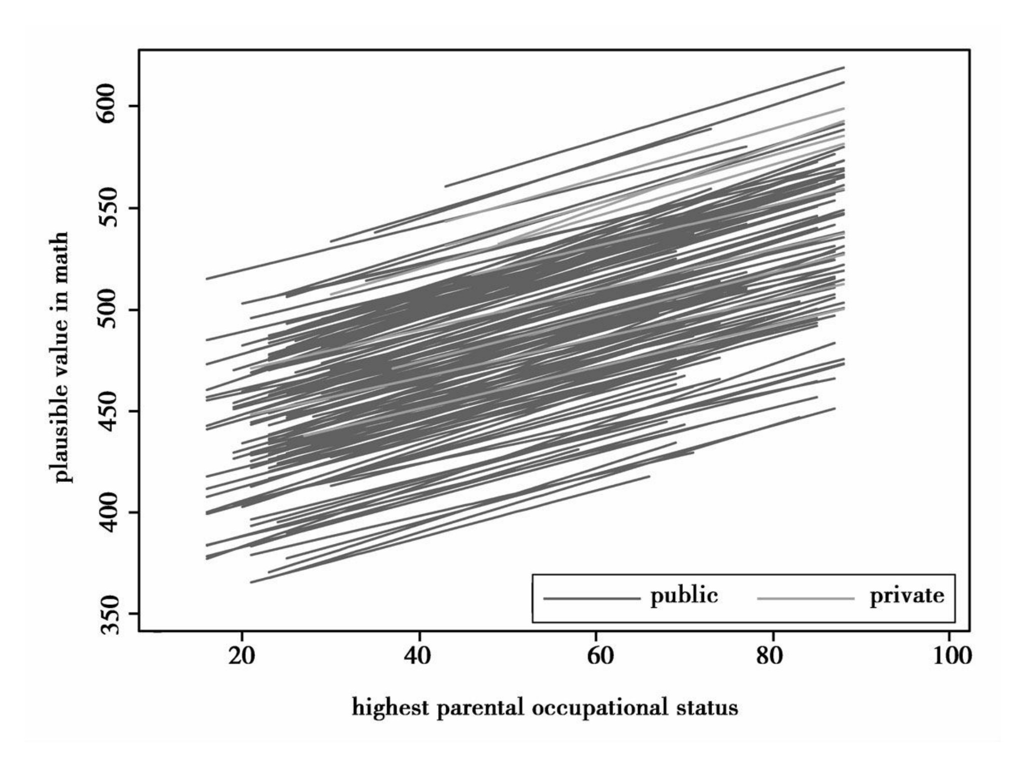


图6-8　多水平回归模型图

6.4 M*plus* 代码与结果呈现

某些统计软件如Stata或R中，多水平回归模型代码基于复合模型（本例中即方程6.16）构建，在M*plus* 中则由不同水平的联立方程组（本例中即方程6.13、6.14、6.15）构建。对于上述多水平回归模型，对应的M*plus* 代码见框6-1。

框6-1　典型的多水平回归模型

```
Title: Multilevel Model for PISA
Data:
  FILE IS pisa. dat;
Variable:
  Names＝
    schoolid hisei pv1math sector;
  MISSING＝all（－9999）;
  USEVARIABLES＝hisei pv1math sector;
  WITHIN＝hisei;
  BETWEEN＝sector;
  CLUSTER＝schoolid;
! DEFINE: CENTER hisei（GROUPMEAN）;     ! 组均值中心化
! DEFINE: CENTER hisei（GRANDMEAN）;     ! 总体均值中心化
ANALYSIS: TYPE＝TWOLEVEL RANDOM;
  ESTIMATOR＝ML;
MODEL:
  %WITHIN%
```

```
  s | pv1math ON hisei
%BETWEEN%
  pv1math s ON sector;
  pv1math WITH s;       ! 水平-2残差协方差
```

上述代码中，数据定义、变量声明部分与一般的模型无异，但作为多水平回归模型，有几处特殊的地方需要注意：

①VARIABLE部分，需要声明区分不同组群的CLUSTER变量，本例中即学校编码schoolid。

②对于多水平回归模型，需要在VARIABLE部分，指明变量所属的水平。具体到两水平模型而言，即指明变量属于WITHIN还是BETWEEN层面。不过，在M*plus* 代码书写规范上，对于不同水平都要用到的变量，即所谓“组内/组间两用”（within-between status）变量，不要声明其所属层次。[(10)]

③ANALYSIS部分，对于两水平模型，设置TYPE＝TWOLEVEL。如果模型有随机斜率，需要增加RANDOM选项。

④MODEL部分，各水平的方程分别设定，WITHIN部分即水平-1方程，BETWEEN为水平-2方程。

⑤“pv1math s ON sector”，这是两个方程写在一行（当然也可以分别写），即水平-2针对水平-1截距β_{0j}和斜率β_{1j}的方程。需要注意的是，这里水平-2针对水平-1斜率β_{1j}的方程，等号左手边

的变量（即s）是新命名的，针对水平-1截距β_{0j}的方程，等号左手边变量名使用的仍是水平-1的结果变量名“pv1math”。[(11)]

⑥水平-1中“s | pv1math ON hisei”中“s | ”代表水平-1方程有斜率随机变动。具体执行中，这样的设置会自动生成水平-2变量sector和水平-1变量hisei的跨层交互项。

运行上述代码，即可在output文件得到相应结果，见框6-2。M*plus* 的多水平回归模型输出结果风格与多数统计软件（如HLM）有所不同，如果是初次接触，解读起来可能会有些令人迷惑。最好结合模型的定义进行，或者与熟悉的多水平处理软件或程序（如HLM、Stata等）的输出结果进行比照解读。

框6-2　典型的多水平回归模型输出结果（部分）

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
PV1MATH	5444.667	114.084	47.725	0.000
Between Level				
S ON				
SECTOR	-0.067	0.347	-0.193	0.847
PV1MATH ON				
SECTOR	32.509	24.671	1.318	0.188
PV1MATH WITH				
S	-6.239	6.183	-1.009	0.313
Intercepts				
PV1MATH	421.336	5.211	80.848	0.000
S	1.215	0.082	14.742	0.000
Residual Variances				
PV1MATH	1621.901	445.938	3.637	0.000
S	0.151	0.107	1.411	0.158

省略其他部分，直接对“MODEL RESULTS”进行解读，模型结果被分为“Within Level”和“Between Level”两个部分。由于水平-1的

截距和斜率都是由水平-2因素决定的，或者说其变化都体现在水平-2，方差存在与水平-2。对本例而言，各系数γ_{00}、γ_{01}、γ_{10}、γ_{11}、μ_{0j}、μ_{1j}其实都是水平-2方程的要素。因此，在M*plus* 输出结果中，“Within Level”部分只有水平-1的残差方差，即σ^2，其值为5444.667。

“Between Level”部分，“S ON SECTOR”即方程6.15中sector对β_{1j}的系数，即γ_{11}，从复合模型6.16来看，也就是sector和hisei交互项的系数，其值为－0.067。从结果来看统计上并不显著。

“*P* V1MATH ON SECTOR”即方程6.14中sector对β_{0j}的系数，即γ_{01}，其值为32.509，实际意义也就是私立学校比公立学校学生数学成绩平均分高32.509分。

“PV1MATH WITH S”即β_{1j}与β_{0j}的协方差，即τ_{01}。

“Intercepts”类目下列了“PV1MATH”和“S”两行，分别对应截距与斜率的均值，即γ_{00}和γ_{10}。由于均随机变化，这些值在各学校之间不同，这里给出其均值。具体分别为421.336和1.215。由于存在sector和hisei的交互项，具体解释还要结合交互项的情况进行。

“Residual Variances”类目下，“*P* V1MATH”和“S”两行，则分别对应方程6.14和方程6.15中截距和斜率的误差项，即u_{0j}和u_{1j}。

当然，如果是要利用模型结果发表论文的话，不能直接使用原始结果，需要整理成通常的出版格式汇报表。表6-4即为根据上述结果整理而成。一般要按“固定效应”和“随机效应”的方式分别进行整理。用于评估模型拟合和比较的其他一些信息列在下方。

表6-4　多水平回归结果

	（1） 原始变量	（2） 总中心化	（3） 组中心化
Fixed Effect			
Intercepts（γ_{00}）	421.336*** （5.211）	484.011*** （3.317）	482.625*** （3.738）
sector（γ_{01}）	32.509 （24.671）	29.070* （13.945）	44.156** （15.265）
hisei（γ_{10}）	1.215*** （0.082）		
hiseiXsector（γ_{11}）	−0.067 （0.347）		
grandhisei（γ_{10}）		1.215*** （0.082）	
grandsesXsector（γ_{11}）		−0.067 （0.347）	
grouphisei（γ_{10}）			1.133*** （0.082）
grouphiseiXsector（γ_{11}）			−0.077 （0.350）
Random Effect			
Intercepts（τ_{00}）	1621.901*** （445.938）	1379.343*** （181.311）	1819.481*** （229.516）
hisei（τ_{11}）	0.151 （0.107）	0.150 （0.107）	0.140 （0.106）
Residual（σ^2）	5444.667*** （114.084）	5444.779*** （114.088）	5444.704*** （114.152）
Model Fit			
N	4842	4842	4842
LL	−27876.335	−27876.335	−27894.310
AIC	55768.670	55768.670	55804.619
BIC	55820.551	55820.550	55856.500

括号内为标准误

$p<0.1$，*$p<0.05$，**$p<0.01$，*** $p<0.001$

6.5 “中心化”问题

表6-4中，除了汇报了使用hisei原始变量进行多水平分析的结果，还汇报了对该变量进行中心化（centering）的结果。多水平模型中，中心化也称“对中”或“平减”，是一个常见的处理。也就是将某原始变量减去对应的总均值或组均值，作为新的变量带入模型。

中心位置的选择主要取决于研究目的，并没有普遍使用的原则。中心化一般有两种方式：总均值中心化（grand-mean centering）和组均值中心化（group-mean）。另外，在增长模型中，涉及不同时点的观测数据，可选择具有特定现实意义的时间点作为基点进行定位。

中心化可以让结果便于解释。一般多元回归模型中，人们主要关注斜率。多水平回归模型中，水平-1截距和斜率为宏观方程的因变量，因此水平-1随机回归系数尤其是截距有意义的解释就十分重要，否则水平-1随机系数在组水平变异的分析就不好解释。

当水平-1自变量值取零且截距意义不清楚时，对其进行中心化变换可使得截距有合理解释。比如，若SES取值为200～600，那么β_{0j}应为第j个学校当SES＝0时结果变量的取值。此时β_{0j}并无意义，而且导致截距和斜率之间高度的相关性（Raudenbush＆Bryk，2002：32-35）。此时，可以通过中心化定义或转化SES的测量值让截距变得有意义。如将每一个SES减去样本平均的SES。相当于SES均值的测量离差（deviation from the mean），代表某个体SES在样本均值中的相

对位置。这样，具有样本SES均值的个体中心化SES为0，而截距β_{0j}就代表样本中具有平均SES者的结果变量的期望值。

经过中心化处理，变量测量值变成了相对值，是某个个体测量值与某常数之间的差，代表某个体在组群中的相对位置。因此，需注意测量转换后变量效应的解释。[(12)]

水平-2自变量是否中心化并不关键，一般不涉及稳定性的问题。除非引入了水平-2自变量之间的交互项（Raudenbush＆Bryk，2002：35）。在多水平回归模型中，水平-2自变量一般进行总均值中心化，但与不进行中心化的自然测量模型相比，结果一致。

考虑到场景效应与跨水平交互作用，自变量采取原始形式和采用总均值中心化的结果，可视为相同模式（Hofmann＆Gavin，1998）。

因变量不进行中心化。

中心化的好处还包括：中心化可以去掉随机截距与斜率之间的高相关，以及水平-1、水平-2与跨水平交互作用之间的高相关，从而改善了多重共线性引起的水平-2参数估计问题。中心化有助于解释交互作用中变量的主效应。使得模型更为稳定，系数更具独立性。有助于提高模型运行速度，减少模型估计收敛不好的问题（*P* accagnella，2006）。

中心化也存在一定风险，可能会让一个并不符合研究需要的模型变得更为拟合。尤其是组均值中心化不论是从统计上还是概念上来看，都与其他两种有所不同，中心化方法的选择必须特别小心。总均值中心化只是改变一些参数的值，并不改变模型；而组均值中心化模

型的参数不是自然测量模型参数的简单函数。研究者可能会冒风险去估计一个不同于原分析思想的模型（Hofmann＆Gavin，1998；Paccagnella，2006）。

如果水平-1以组均值中心化，最好同时将各组均值纳入模型，以利于组均值效应的修正（Kreft＆De Leeuw，1998）。对于随机斜率模型，自变量最好不要采用组均值中心化，除非有明确的理论指导或经验依据，表明组均值中心化自变量与结果变量更相关。组中心化值暗示了个体在其所属群体内的相对位置，在对如参照群体、相对剥夺、教师排名对学生表现影响等问题中采用更为合适（Snijders＆Bosker，2011）。

有研究指出，中心化方法的选择不能基于模型的统计依据，而应该基于研究的具体问题。一些经验法则，如果研究兴趣是水平-1自变量对因变量的影响，选择组均值中心化是合适的；如果主要兴趣是场景效应则用总均值中心化；要研究跨水平交互作用，按组均值中心化较好，而研究水平-2变量间的交互作用则按总均值中心化（Enders＆Tofighi，2007）。

如果研究者的兴趣是水平-2自变量对结果变量的递增效果（incremental），用总均值中心化是合适的，因为组均值中心化不能充分控制水平-1变量的影响；如果研究中介效应，用组均值和总均值中心化均可；如果研究水平-2自变量的调节作用，则建议用组均值中心化；如果研究兴趣是组内个体特征对结果变量的影响而不考虑组间的影响，即“青蛙池塘效应（frog pond effect）”，适合采用组均值中心化（Hofmann＆Gavin，1998）。

如果研究者不太关心组织层之间的差异，更关注的是个人层解释变量，不进行中心化处理更好（Kreft et al，1995）。研究者的目的是个人效应时，不对变量进行中心化是很好的选择。如果分析的目的是把组织层效应从个人层特征区别出来，那么选择对预测变量进行组均值中心化是比较好（*P* accagnella，2006）。当总均值中心化和不进行中心化差别不大时，Kreft et al（1995）推荐采用总均值中心化，可以减轻共线性问题。

进行跨层交互作用的分析时，可以使用原始变量或者总中心化处理的变量。不过，由于此时水平-1的斜率包含了组内和组间的关系，因此跨层交互作用可能存在虚假成分。为此，Hoffmann & Gavin（1998）建议对于此类问题，水平-1预测变量使用组中心化，同时在水平-2增加组均值作为控制变量。实际建模时可以两种方法同时估计，进行比较（廖卉、庄瑗嘉，2012）。对水平-1预测变量采取组中心化处理的优势是，系数估计只包含组内效果，而且跨层交互和组间交互作用（采取总中心化）互相独立（Enders＆Tofighi，2007）。近来从多水平结构方程模型角度进行中介分析的文献指出，在基于多水平回归模型的中介效应分析中，为了避免组内和组间效应的混淆，需要对水平-1预测变量进行组中心化处理（Krull＆MacKinnon，2001；*P* reacher et al，2011；Zhang et al，2009）（参见第9章相关内容。）

总之，研究者需要根据研究目的审慎地考虑选择哪种中心化模型，解释结果时也需要考虑中心化的影响。有关截距和截距标准误的解释，与截距相关的协方差的解释，要特别小心（Raudenbush＆Bryk，2002：32-35）。

在M*plus* 中，如果数据中的变量未提前进行中心化，可以在M*plus*中通过DEDINE命令实现，具体参见框6-1中DEFINE部分（源码中被注释掉）。

6.6 跨水平交互作用

跨水平交互作用（cross-level interactions）或微观-宏观交互作用（micro-macro interactions）是多水平分析中的一个重要问题。跨水平交互项具有重要的理论与实际意义，是多水平建模需要关注的一个重点。

如果水平-1解释变量X_{ij}对于结果变量Y_{ij}的作用在各组之间有显著变异，即β_{1j}非固定，那么就需要进一步分析哪些情境变量会影响水平-1解释变量与结果变量之间的关系。

如果在水平-2方程中，某情境变量W_{1j}对水平-1斜率系数β_{1j}效应统计上显著，表明水平-1解释变量与结果变量之间的关系取决于水平-2的解释变量W_{1j}，也就是受到W_{1j}的影响和调节。像在一般多元回归模型中一样，如果跨层交互项统计上显著，主效应也必须同时保留在模型之中，并将相应的效应计算在内。根据目前数据，跨层交互项并不显著。不过这里以此为例，简单交代一下交互作用的估计。忽略或者说平均掉随机效应，因变量均值可以表达为：

$$\overline{\text{pv1math}_{ij}} = \gamma_{00} + \gamma_{10}\text{hisei}_{ij} + \gamma_{01}\text{sector}_j + \gamma_{11}\text{sector}_j \times \text{hisei}_{ij} \qquad (6.17)$$

如果想得到hisei对因变量pv1math均值的偏效应（partial effect），可计算有关hisei的斜率方程：

$$\frac{\partial\ \overline{\text{pv1math}_{ij}}}{\partial \text{hisei}_{ij}} = \gamma_{10} + \gamma_{11}\text{sector}_j \tag{6.18}$$

本例中sector只取0、1两个变量，所以其值在1.215＋（－0.067）×0＝1.215和1.215＋（－0.067）×1＝1.148之间变化。

$$\text{Var}(\hat{\gamma}_{10} + \hat{\gamma}_{11}\text{sector}_j) =$$
$$\text{Var}(\hat{\gamma}_{10}) + 2 \times \text{sector}_j \times \text{Cov}(\hat{\gamma}_{10},\hat{\gamma}_{11}) + \text{sector}_j^2 \times \text{Var}(\hat{\gamma}_{11}) \tag{6.19}$$

根据系数方差协方差矩阵估计结果，$\text{Var}(\hat{\gamma}_{10})$ ＝0.006655，Cov $\text{Cov}(\hat{\gamma}_{10},\hat{\gamma}_{11})$ ＝－0.006655，Var（γ_{11}）＝0.11900535。带入下式即可得95%置信区间：

$$\hat{\gamma}_{10} + \hat{\gamma}_{11}\text{sector}_j \pm 1.96 \times \sqrt{\text{Var}(\hat{\gamma}_{10} + \hat{\gamma}_{11}\text{sector}_j)} \tag{6.20}$$

M*plus* 可以输出调节效应图，见框6-3。代码解释见相应注释。

框6-3　跨层交互作用可视化

```
TITLE: Verfying Cross-level Interactions for PISA
DATA:
FILE IS pisa. dat;
VARIABLE:
  NAMES＝schoolid hisei pv1math sector;
  MISSING＝all（－9999）;
  USEVARIABLES＝pv1math hisei sector;
  CLUSTER＝schoolid;
  WITHIN＝hisei;
```

```
  BETWEEN＝sector;
ANALYSIS: TYPE＝TWOLEVEL RANDOM;
  ESTIMATOR＝ML;
MODEL:
  %WITHIN%
  s|pv1math ON hisei;
  %BETWEEN%
  pv1math ON sector;
  ［s］（gam0）；          ！水平-2第一个方程，Beta__1j取均值即截距Gamma__10
  s ON sector（gam1）；   ！水平-2第二个方程，sector和Beta__1j的斜率系数Gamma__11
  pv1math WITH s;
MODEL CONSTRAINT:
  PLOT（ylow yhigh）；
  LOOP（XVAL，－2，2，0.5）；        ！设定演示的取值范围
  ylow＝（gam0＋gam1*0）*XVAL;   ！上述偏微分公式，sector 0 1变化，连续变量取＋－SD
  yhigh＝（gam0＋gam1*1）*XVAL;   ！未含误差区间；不考虑随机效应
PLOT:
  TYPE＝plot2;
```

代码执行输出图形后，点击PLOT菜单下View *P* lots按钮，M*plus*可以输出跨水平调节作用的图形。使用程序自带相应图形编辑功能（图形上单击鼠标右键可以看到）整理后，输出图形见图6-9。

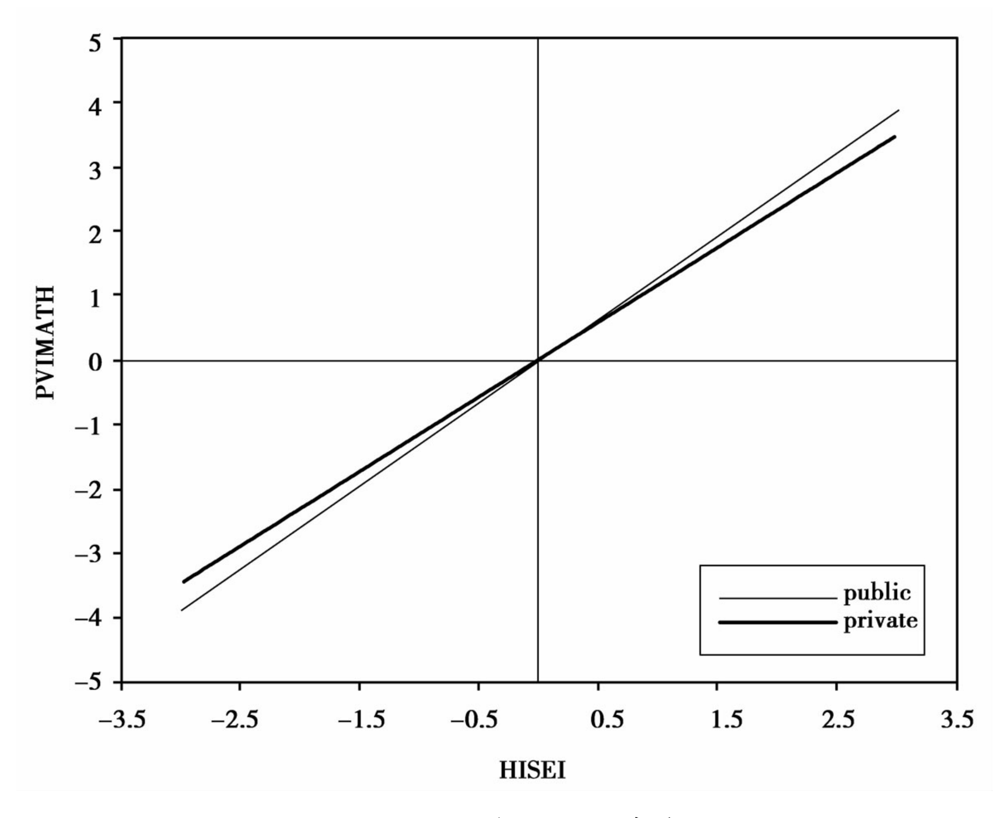


图6-9　*Mplus* 跨水平调节作用图

6.7　常见子模型

6.7.1　随机效应单因素方差分析

随机效应单因素方差分析（one-way ANOVA with random effects）是最简单的多水平回归模型，又称为空模型（empty model）或完全无条件模型（fully unconditional model），模型中不包括任何预测变量：

水平-1：

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij} \tag{6.21}$$

水平-2：

$$\beta_{0j} = \gamma_{00} + \mu_{0j} \tag{6.22}$$

将式6.22代入式6.21，得复合模型：

$$Y_{ij} = \gamma_{00} + \mu_{0j} + \varepsilon_{ij} \tag{6.23}$$

其中，ε_{ij}～N（0，σ^2），μ_{0j}～N（0，τ_{00}）且Cov（ε_{ij}，μ_{0j}）＝0。

模型对应的代码见框6-4。

框6-4　带有随机效应的单因素方差分析代码

```
Title: One-Way ANOVA with Random Effects for PISA
Data:
FILE IS pisa. dat;
Variable:
  Names＝
  schoolid hisei pv1math sector;
  MISSING＝all（－9999）;
  USEVARIABLES＝pv1math;
  CLUSTER＝schoolid;
  WITHIN＝;   BETWEEN＝;
ANALYSIS: TYPE＝TWOLEVEL;
  ESTIMATOR＝ML;
```

```
MODEL:
  %WITHIN%
  pv1math;
  %BETWEEN%
  pv1math;
```

运行上述代码，输出结果见框6-5。

框6-5　带有随机效应的单因素方差分析输出结果（部分）

```
SUMMARY OF DATA
Estimated Intraclass Correlations for the Y Variables
Intraclass
Variable    Correlation
PV1MATH         0.250
MODEL RESULTS
                                                              Two-Tailed
                      Estimate        S.E.    Est./S.E.       P-Value
Within Level
Variances
    PV1MATH           5753.248     133.429       43.118          0.000
Between Level
Means
    PV1MATH            485.286       3.723      130.347          0.000
Variances
    PV1MATH           1922.240     214.737        8.952          0.000
```

空模型的一个重要用途是可以计算PISA数据的ICC，从而判断是否需要进行多水平分析。上述结果中已输出了ICC＝0.250。当然也可以自己计算，根据上述结果：

$$\begin{aligned} Var(\mu_{0j}) &= 1922.240 \\ Var(\varepsilon_{ij}) &= 5753.248 \end{aligned} \tag{6.24}$$

那么，ICC：

$$\rho = \frac{\tau_{00}}{\tau_{00} + \sigma^2} = \frac{1922.240}{1922.240 + 5753.248} \approx 0.250 \qquad (6.25)$$

也就是说，学生数学成绩呈现较强的聚类效应，1／4的变差可以归结为学校之间的差异。该值远远大于0.059，因此有必要进行多水平分析。

6.7.2　以均值为结果的回归

以分组特征预测各组均值，即为以均值为结果的多水平回归模型（means-as-outcomes regression），其特点为水平-1为空模型，水平-2有自变量：

水平-1：

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij} \qquad (6.26)$$

水平-2：

$$\beta_{0j} = \gamma_{00} + \gamma_{01} W_j + \mu_{0j} \qquad (6.27)$$

将式6.27代入式6.26，得复合模型：

$$Y_{ij} = \gamma_{00} + \gamma_{01} W_j + \mu_{0j} + \varepsilon_{ij} \qquad (6.28)$$

其中，ε_{ij}～N（0，σ^2），μ_{0j}～N（0，τ_{00}）且Cov（ε_{ij}，μ_{0j}）＝0。

输出结果显示省略，模型对应的代码见框6-6。

框6-6　以均值为结果的多水平回归代码

```
Title: Means-as-Outcomes Regression for PISA
Data: FILE IS pisa. dat;
Variable:
  Names＝schoolid hisei pv1math sector;
  MISSING＝all（－9999）;
  USEVARIABLES＝pv1math sector;
  CLUSTER＝schoolid;
  WITHIN＝;
  BETWEEN＝sector;
ANALYSIS:
  TYPE＝TWOLEVEL;
  ESTIMATOR＝ML;
MODEL:
  %WITHIN%
  pv1math;
  %BETWEEN%
  pv1math ON sector;
```

虽然该模型仍然十分简单，但与整体回归和聚合回归相比，这类模型作为随机截距模型，提供了固定效应无偏和有效估计，同时还提供了更适当的标准误估计（Raudenbush＆Bryk，2002：107-109）。

6.7.3 随机效应单因素协方差分析

在标准的多水平回归模型方程6.13、6.15和6.16中，水平-2系数 γ_{01} 、γ_{11} 和 μ_{1j} 设为0，即截距不等而斜率相同，水平-1含自变量作为协变量，可得所谓带有随机效应的单因素协方差分析（one-way ANCOVA with random effects）模型：

水平-1：

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij} \quad (6.29)$$

水平-2：

$$\beta_{0j} = \gamma_{00} + \mu_{0j} \quad (6.30)$$

$$\beta_{1j} = \gamma_{10} \quad (6.31)$$

将式6.30、式6.31代入式6.29，得复合模型：

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \mu_{0j} + \varepsilon_{ij} \quad (6.32)$$

其中，ε_{ij}～N（0，σ^2），μ_{0j}～N（0，τ_{00}）且Cov（ε_{ij}，μ_{0j}）＝0。

上式中也可以引入水平-2协变量，如 γ_{01} 不能省略，组合模型将呈现以下形式：

$$Y_{ij} = \gamma_{00} + \gamma_{01}W_j + \gamma_{10}X_{ij} + \mu_{0j} + \varepsilon_{ij} \quad (6.33)$$

注意：该模型与传统协方差分析的不同之处在于，分组效应μ_{0j}是随机而非固定的；如果自变量中心化需要以总均值为基准。

输出结果显示省略，模型对应的代码见框6-7。

框6-7　随机效应单因素协方差分析代码

```
Title: One-Way ANCOVA with Random Effects for PISA
Data:
FILE IS pisa. dat;
Variable:
  Names＝schoolid hisei pv1math sector;
  MISSING＝all（－9999）;
  USEVARIABLES＝pv1math hisei;
  CLUSTER＝schoolid;
  WITHIN＝hisei;
  BETWEEN＝;
  DEFINE: CENTER hisei（GRANDMEAN）
ANALYSIS:
  TYPE＝TWOLEVEL;
  ESTIMATOR＝ML;
MODEL:
  %WITHIN%
  pv1math ON hisei;
  %BETWEEN%
  pv1math;
```

6.7.4　随机系数回归模型

水平-2为空模型，但水平-1含自变量，即得随机系数模型（random-coefficients regression model）。该模型允许水平-1截距和斜率在水平-2单位之间随机变化。

水平-1：

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + \varepsilon_{ij} \quad (6.34)$$

水平-2：

$$\beta_{0j} = \gamma_{00} + \mu_{0j} \quad (6.35)$$

$$\beta_{1j} = \gamma_{10} + \mu_{1j} \quad (6.36)$$

将式6.35、式6.36代入式6.34，得复合模型：

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \mu_{0j} + \mu_{1j} X_{ij} + \varepsilon_{ij} \quad (6.37)$$

其中，

$$\varepsilon_{ij} \sim \mathrm{N}(0, \sigma^2)$$

$$\begin{pmatrix} \mu_{0j} \\ \mu_{1j} \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix} \right)$$

输出结果显示省略，模型对应的代码见框6-8。

框6-8　随机系数回归模型代码

```
Title: Random-Coefficients Regression Model for PISA
Data:
FILE IS pisa. dat;
Variable:
  Names＝
  schoolid hisei pv1math sector;
  MISSING＝all　（－9999）;
  USEVARIABLES＝pv1math hisei;
  CLUSTER＝schoolid;
  WITHIN＝hisei;
  BETWEEN＝;
ANALYSIS: TYPE＝TWOLEVEL RANDOM;
  ESTIMATOR＝ML;
MODEL:
  %WITHIN%
  s | pv1math ON hisei;
  %BETWEEN%
  pv1math;
  s;
  pv1math with s;
```

6.7.5　以截距和斜率为结果的模型

以截距和斜率为结果（intercepts-and slopes-as-outcomes）的模型也称为随机变动系数模型（random varying coefficient model），指水平-1回归方程中截距和斜率作为水平-2回归方程因变量。与随机系数模型不同之处在于，水平-2为非空模型。

一般来讲，水平-1有n个自变量，加上截距项，水平-2要构建$n+1$个方程分别各自对应。

以截距和斜率为结果属于典型的多水平回归模型，前文已经介绍模型设定（见方程6.13—6.16）、代码及结果，这里不再赘述。

6.7.6　非随机变动斜率模型

当水平-2自变量W_j充分预测水平-1斜率β_{1j}的变化（控制W_j后μ_{1j}接近于0），为了提高统计上有效和计算稳定性，可设μ_{1j}为0，从而得到非随机变动斜率模型（model with nonrandomly varying slopes），其特点是水平-2斜率方程为固定效应，水平-1斜率虽然在各组之间变化，但变化非随机。

水平-1：

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + \varepsilon_{ij} \quad (6.38)$$

水平-2：

$$\beta_{0j} = \gamma_{00} + \gamma_{01} W_j + \mu_{0j} \quad (6.39)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} W_j \quad (6.40)$$

将式6.39、式6.40代入式6.38，得复合模型：

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} W_j + \gamma_{11} W_j X_{ij} + \mu_{0j} + \varepsilon_{ij} \quad (6.41)$$

其中，$\varepsilon_{ij} \sim N(0, \sigma^2)$，$\mu_{0j} \sim N(0, \tau_{00})$ 且$Cov(\varepsilon_{ij}, \mu_{0j}) = 0$。

输出结果显示省略，模型对应的代码见框6-9。

框6-9　非随机变动斜率模型代码

```
Title: Nonrandomly Varying Slopes Model for PISA
Data:
FILE IS pisa. dat;
Variable:
  Names＝schoolid hisei pv1math sector;
  MISSING＝all （－9999）;
  USEVARIABLES＝pv1math hisei sector;
  CLUSTER＝schoolid;
  WITHIN＝hisei;
  BETWEEN＝sector;
ANALYSIS: TYPE＝TWOLEVEL RANDOM;
```

```
  ESTIMATOR＝ML;
MODEL:
  %WITHIN%
  s | pv1math ON hisei;
  %BETWEEN%
  pv1math ON sector;
  s ON sector;
  s@0;        ! 设定斜率方差为零等价于设定误差项为零
  pv1math with s;
```

其中，“s@0”代表将斜率方差固定为0，也就是不再随机变动，只是根据sector变量的变化而有变化。

6.7.7　小结

为了便于比较，将上述各模型总结见表6-5。

表6-5　两水平模型各子模型

	水平-1	水平-2	
		截距	斜率
1.随机效应单因素方差分析	$Y_{ij} = \beta_{0j} + \varepsilon_{ij}$	$\beta_{0j} = \gamma_{00} + \mu_{0j}$	
2.以均值为结果的回归	$Y_{ij} = \beta_{0j} + \varepsilon_{ij}$	$\beta_{0j} = \gamma_{00} + \gamma_{01} W_j + \mu_{0j}$	
3.随机效应单因素协方差分析	$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \overline{X}_{..}) + \varepsilon_{ij}$	$\beta_{0j} = \gamma_{00} + \mu_{0j}$	$\beta_{1j} = \gamma_{10}$
4.随机系数回归模型	$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + \varepsilon_{ij}$	$\beta_{0j} = \gamma_{00} + \mu_{0j}$	$\beta_{1j} = \gamma_{10} + \mu_{1j}$
5.以截距和斜率为结果的模型	$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + \varepsilon_{ij}$	$\beta_{0j} = \gamma_{00} + \gamma_{01} W_j + \mu_{0j}$	$\beta_{1j} = \gamma_{10} + \gamma_{11} W_j + \mu_{1j}$
6.非随机变动斜率模型	$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + \varepsilon_{ij}$	$\beta_{0j} = \gamma_{00} + \gamma_{01} W_j + \mu_{0j}$	$\beta_{1j} = \gamma_{10} + \gamma_{11} W_j$

其中，模型1、2、3、6属于随机截距模型；模型4、5属于随机变化斜率模型。混合模型、情境模型、Cronbach模型等都可在多水平回归模型框架下来表示，其均可视为多水平回归模型的特殊形式。

在选择模型时候，需要考虑：水平-2不同组群截距是否相同？能否用水平-2协变量预测截距之间的差异？以及，斜率是否存在水平-2的变异？能否用水平-2预测变量解释斜率之间的差异？

一般情况下，水平-1包含多个解释变量，可以根据理论假设和实际情况，将斜率设定为随机、非随机变动以及固定形式的任何组合。截距亦可。

按照出版格式，将上述6个模型的结果整理到表6-6中。

表6-6　多水平分析各模型

	(1)	(2)	(3)	(4)	(5)	(6)
Fixed Effect						
Intercepts (γ_{00})	485.286*** (3.723)	482.660*** (3.761)	485.867*** (3.251)	422.792*** (5.122)	421.336*** (5.211)	421.204*** (5.263)
sector (γ_{01})		44.084** (13.931)			32.509 (24.671)	35.251+ (19.643)
hisei (γ_{10})			1.222*** (0.073)	1.219*** (0.080)	1.215*** (0.082)	1.220*** (0.083)
hiseiXsector (γ_{11})					−0.067 (0.347)	−0.102 (0.346)
Random Effect						
Intercepts(τ_{00})	1922.240*** (214.737)	1813.359*** (204.077)	1429.185*** (186.954)	1651.775*** (447.244)	1621.901*** (445.938)	1350.436*** (342.899)
hisei (τ_{11})				0.154 (0.107)	0.151 (0.107)	
Residual (σ^2)	5753.248*** (133.429)	5753.070*** (133.419)	5477.431*** (113.158)	5444.662*** (114.082)	5444.667*** (114.084)	5476.861*** (125.876)
Model Fit						
N	4842	4842	4842	4842	4842	4842
LL	−28015.996	−28011.922	−27879.976	−27878.507	−27876.335	−27877.606
AIC	56037.991	56031.844	55767.952	55769.015	55768.670	55769.211
BIC	56057.447	56057.784	55793.892	55807.925	55820.551	55814.607

括号内为标准误；模型3中hisei采取了总平均化形式

$^{+}p<0.1$，$p<0.05$，$^{**}p<0.01$，$^{***}p<0.001$.

6.8　模型估计

对于多水平回归模型的估计问题，首先要区分模型、模型估计和算法的概念。模型是不同变量之间关系的方程，比如两水平回归模型。模型估计是将某种统计原理应用于具体模型的结果，比如使用最大似然法去估计模型的参数。算法（algorithm）用于模型和统计技术，指的是具体的运算技术，如EM（Kreft＆De Leeuw，1998；Muthén＆Muthén，2012）。

M*plus* 中提供了基于贝叶斯（Bayesian）和频次（frequentist）的不同方式。贝叶斯分析使用马尔科夫链蒙特卡洛（Markov chain Monte Carlo）算法。频次方法使用最大似然法（maximum likelihood，ML）、加权最小二乘法（generalized or weighted least squares，GLS）等。

最大似然估计包括完全信息最大似然估计（full information maximum likelihood；FIML）和受限最大似然估计（restricted or residual maximum likelihood；REML）。FIML在似然函数中包括了回归系数和方差成分。REML似然函数中仅包括方差成分，回归系数在第二步进行估计。某些软件如SAS中REML是默认的估计方法。要注意的是，M*plus* 中没有提供REML，M*plus* 中的ML即FML（FIML）。[(13)]

对于似然法求解，M*plus* 提供了拟牛顿算法（Quasi-Newton）、Fisher评分算法（Fisher Scoring）、牛顿-拉夫逊方法（Newton-Raphson ）以及期望最大化算法（Expectation Maximization，EM）等最优化算法。比如，EM算法就是一种处理当数据中存在缺省值或随机参数时计算最大似然估计值的迭代方法。最大期望算法经过两个步

骤交替进行计算：第一步是计算期望（E），利用现有估计值，计算其最大似然估计值；第二步是最大化（M），最大化在E步上求得的最大似然值来计算参数的值。M步上找到的参数估计值被用于下一个E步计算中，这个过程不断交替进行。[14]

贝叶斯估计将任何有关参数的先验信息与数据中包含的信息结合起来，生成一个后验分布。然后使用MCMC算法不断迭代，进行特定抽样（Gibbs sampling and Metropolis-Hastings sampling）形成随机样本，不断逼近后验分布。由于可以利用某组样本和整体样本数据计算最佳加权平均参数估计值，相当于借用全部样本信息支持样本量较小组群的参数估计，对于某些组群出现稀少数据（sparse data）更为适用，也适合于某些组样本量较小的情况。

M*plus* 进行多水平建模的一个优势是，不仅可以处理因变量为连续变量的一般多水平回归模型，而且可以允许因变量是二分变量、有序或无序类别变量、截取变量、计数变量等多种类型，运算速度也相对较快，这是目前其他一些统计软件难以比肩的。对于两水平模型，M*plus* 提供的主要估计方法见表6-7。[15]

表6-7　多水平回归模型估计方法及适用情况

估计方法	适用情况
FIML	允许连续、删节、二分、多分类、定序、计数等变量形式，或这些变量类型的组合；允许随机截距和斜率，允许有缺失值。对于历时数据，允许个体观测次数不同及时变协变量带随机斜率。对于分布非正态，使用MLR可以获得稳健标准误，并输出模型拟合卡方检验。

WLS	允许连续、二分、定序变量及这些变量类型组合；允许随机截距；允许缺失值。
MUML	仅限于连续变量、随机截距，不允许数据有缺失值。
BAYES	允许连续、类别及这些变量类型组合；随机截距和斜率，允许缺失值。

M*plus* 中，估计方法需要在Analysis部分指定，比如：

框6-10　指定估计方法和算法

```
Analysis:
  TYPE＝TWOLEVEL RANDOM;
  ESTIMATOR＝MLR;                    ！使用带有稳健标准误的最大似然估计
  ALGORITHM＝INTEGRATION;            ！采取数值积分法优化算法，另EM、EMA、FS等
  INTEGRATION＝MONTECARLO;           ！采用MONTECARLO积分，另STANDARD等
  PROCESSORS＝4;                     ！四核处理器
```

大多数情况下，各类算法都需要进行迭代运算。估计开始，依据一定规则产生初始值，经过一次迭代产生的参数估计值作为下一次迭代的初始值。如此循环，直到参数估计趋于稳定位为止。收敛的标准可以通过设置容限度（tolerance）或收敛标准（convergence criterion）来控制，一般统计软件也有默认值（M*plus* 默认为0.001）。

实际分析中，经常会遇到模型估计迭代难以收敛的问题。收敛问题往往由变量测量尺度悬殊、初始值问题，或者模型不可识别等引起。尤其是混合模型、多水平模型以及带有随机效应（方差较小）的模型的情况下，更容易出现不收敛的问题（参见Muthén＆Muthén，2012：467）。

为了避免无休无止的循环，统计软件也会设置默认的最大迭代次数（M*plus* 默认为50000次）。当遇到难以收敛的问题时，如果尚未达到默认的迭代次数或收敛标准，可以加大迭代次数和放宽收敛标准（比如MCONVERGENCE＝0.01；ITERATIONS＝30000；）。[(16)]然而，若迭代不收敛是因为模型设定不正确、数据自身问题（如样本量太少、数据有异常值）或者模型中的变量测量尺度差异过大等问题，则需要进行相应的调整后再进行估计。一些较复杂的模型，需要设定合适的初始值才能进行估计。

6.9　模型拟合

对数似然值（log likelihood）可以用来评估模型的拟合程度。M*plus* 输出结果中，提供了log likelihood值。该值越大（绝对值越小），表示模型对数据拟合越好。

一些处理多水平模型的软件，输出－2LL，即－2×loglikelihood，该值越小越好。有的软件（如HLM）还提供了偏差统计（deviance statistic），即－2LL。

M*plus* 也为模型提供了多种信息标准测量（information Criteria）：AIC（Akaike）、BIC（Bayesian）、Adjusted

BIC（Sample-Size Adjusted BIC）。这些值也是越小越好。

$$AIC = -2LL + 2d \quad (6.42)$$

$$BIC = -2LL + d \times \ln(n) \quad (6.43)$$

$$BIC' = -2LL + d \times \ln\left(\frac{n+2}{24}\right) \quad (6.44)$$

对于非嵌套模型，不能使用LR检验（或Deviance Test，Chi-Square Difference test）。此时可以使用信息标准统计量，如AIC、BIC等。实际上这些指标也可以用来对嵌套模型进行比较。这些指标越小，表明模型拟合数据越好。

这些指标可以在相对意义上用来进行基于相同数据的模型比较，但在绝对意义上并不能说明单一模型本身是否拟合数据。

6.9.1 Raudenbush＆Bryk拟合指数

对于一般的多元线性回归模型，R^2 代表解释的方差所占比例（explained proportion of variance），可以衡量模型中自变量对因变量的解释力。但在多水平回归模型中，则没有这么直观，很难给出一个简单适用的R^2 。

Raudenbush＆Bryk（2002）的方法是：水平-1所解释的方差比例，即将自变量X纳入水平-1模型，水平-1削减方差比例，或“解释方差比例”。可以通过比较拟合模型的残差 σ^2 估计值和基准模型或参照模型的 σ^2 估计值得到。水平-1参照模型一般选取带随机效应的ANOVA模型（即空模型）。

$$水平\text{-}1方差解释比例=\frac{\hat{\sigma}^2(带随机效应的ANOVA模型)-\hat{\sigma}^2(拟合模型)}{\hat{\sigma}^2(带随机效应的ANOVA模型)} \quad (6.45)$$

如表6.6中，带有随机效应单因素协方差分析（3）与空模型（1）相比，水平-1增加了hisei变量，模型中 σ^2 由5753.248减少为5477.431，（5753.248－5477.431）／5753.248≈0.048，这意味着该变量增设，使得校内方差减少了4.8％，也就是说社会经济地位可以解释学生层次数学成绩方差的4.8％。

水平-2各 β_q 所解释方差的比例，即将自变量W纳入水平-2模型来解释某一 β_{qj} 时，这一水平-1随机系数（包括截距与斜率）的方差削减比例或解释方差比例。该指标通过比较由拟合模型的残差 τ_{qq} 估计值和某一参照模型的 τ_{qq} 估计值计算而得。水平-2参照模型一般选取随机系数回归模型。

$$\beta_{qj}方差解释比例=\frac{\hat{\tau}^2(随机系数回归模型)-\hat{\tau}^2(拟合模型)}{\hat{\tau}^2(随机系数回归模型)} \quad (6.46)$$

如截距和斜率为结果的模型（5）与随机系数回归（4）模型，水平-2增加了学校类型变量， τ_{00} 由1651.775减少为1621.901，相对变化量为（1651.775－1621.901）／1651.775≈0.018，意味着学校类型解释了学校之间数学成绩实际差异的1.8％。

需要注意的是，水平-2参数解释的方差，是以水平-1模型设置不变为条件的。因此，需要先建立水平-1模型，然后再纳入水平-2预测变量（Raudenbush＆Bryk，2002）。

6.9.2 Snijders＆Bosker拟合指数

Snijders＆Bosker（1999，2012）认为，把随机效应模型估计的方差成分 σ^2 和 τ^2 作为R^2 的替代存在问题：存在多个方差成分导致数个R^2 值，而且增加变量有时会增加而非减少方差成分甚至出现负的R^2 。他们主张仿照一般多元回归模型“削减误差比例”（proportional reduction of error）的方式，将多水平回归模型R_1^2和R_2^2 定义为预测水平-1结果和组均值时的削减误差比例。

对于一般多元线性回归模型，以一种与R^2 等价的方式，构建预测值均方误差消减比例（proportional reduction of the mean squared error of prediction）：

$$R^2 = \frac{\mathrm{Var}(Y_i) - \mathrm{Var}(Y_i - \sum_h \beta_h X_{hi})}{\mathrm{Var}(Y_i)} = 1 - \frac{\mathrm{Var}\left(Y_i - \sum_h \beta_h X_{hi}\right)}{\mathrm{Var}(Y_i)} \tag{6.47}$$

在多水平框架下，需要区分预测个体结果误差削减比例和预测组群均值削减误差比例。暂不计随机斜率，水平-1解释方差比例定义为均方预测误差：

$$R_1^2 = 1 - \frac{\mathrm{Var}\left(Y_{ij} - \sum_h \gamma_h X_{hij}\right)}{\mathrm{Var}(Y_{ij})} \tag{6.48}$$

其中，

$$\mathrm{Var}\left(Y_{ij} - \sum_h \gamma_h X_{hij}\right) = \sigma^2 + \tau_0^2 \tag{6.49}$$

对于结果变量pv1math，方差即Var（Y_{ij} ）可通过空模型（1）中的组内方差$\hat{\sigma}^2$ 和组间方差$\hat{\tau}_0^2$ 加总获得，其值为1922.240＋5753.248＝7675.488。当然也可通过在空模型代码OUTPUT模块中增加SAMPSTAT选项或在其他统计软件中直接计算（不同方式可能略有出入）。水平-1纳入解释变量（3）hisei后，$\hat{\sigma}^2+\hat{\tau}_0^2$ ＝1429.185＋5477.431＝6906.616。因此，R^2 ＝1－6906.616／7675.488≈10.0％。

对于水平-2，方差解释比例也可定义为均方预测误差削减比例。

$$R_2^2 = 1 - \frac{\mathrm{Var}\left(\bar{Y}_{.j} - \sum_h \gamma_h \bar{X}_{h.j}\right)}{\mathrm{Var}(\bar{Y}_{.j})} \tag{6.50}$$

其中，

$$\mathrm{Var}\left(\bar{Y}_{.j} - \sum_h \gamma_h \bar{X}_{h.j}\right) = \frac{\sigma^2}{n} + \tau_0^2 \tag{6.51}$$

其中，n 为样本量。当各组样本量不同时，可以取均值。对于PISA数据，平均组样本量为32。对原始数据按照学校进行聚合后，可方便计算结果变量pv1math组群均值方差即Var（$\bar{Y}_{.j}$ ）为2132.134。水平-2纳入解释变量sector（2）后，$\hat{\sigma}^2/n+\hat{\tau}_0{}^2$ ＝5753.070／32＋1813.359＝1993.1424。因此，R_2^2 ＝1－1993.1424／2132.134≈6.5％。

随机截距模型中，估计R_1^2、R_2^2 非常容易，而在随机斜率模型则比较麻烦。但一般来说，纳入随机斜率后，R_1^2、R_2^2 一般变化不大。最简单的方法是，忽略随机斜率，以相同的固定部分，重新估计模型。即利用随机截距模型，计算R_1^2、R_2^2 。当组样本量较大时，$\hat{\sigma}^2/n$ 趋近于0，两种方法估计的拟合优度值就很接近。

6.10　模型比较

当存在嵌套关系的模型之间进行比较时，如果模型使用的都是ML估计，可以直接使用似然比检验（likelihood ratio test）或离差检验（deviance test）。模型比较的检验可作为多参数Wald检验的替代方式，可以对加入多类别变量转化的多个虚拟变量前后的模型差异进行检验。

$$D_0 - D_1 = -2\ln\left(\frac{L_0}{L_1}\right) = -2(LL_0 - LL_1) \qquad (6.52)$$

比如，对于表6.6随机系数模型（4）和截距与斜率作为结果的模型（5）（均使用ML估计），LL 分别为－27878.507和－27876.335，自由参数分别为6和8。计算两模型离差值D_0-D_1 ＝－2［（－27878.507）－（－27876.335）］＝4.344，df ＝8－6＝2。

借助卡方分布函数（如Stata：chi2tail（2，4.344）＝0.1139；Excel：CHISQ. DIST. RT（4.344，2）＝0.1139）或查卡方分布表，可以计算p值或临界值，从而判断是否存在显著差异。从结果来看两个模型差异不显著，水平-2增加的解释变量没有显著改善模型拟合。

需要注意，M*plus* 中，根据模型估计的方法不同，嵌套模型差异检验方法也要有所不同：

- 对于结果有卡方值输出的模型，可直接进行卡方差异检验（chi-square difference testing）。当在M*plus* 中执行以MLR、MLM或WLSM估计所得 x^2 差异的检验时，需要用S-B校正法（Satorra-Bentler scaling correction）对卡方进行调整（Muthén＆Muthén，2005；Satorra，2000；Satorra＆Bentler，2001）。不过，MLM不适用多水平回归模型，WLSM估计多水平回归模型不能含

随机斜率，所以对于多水平回归模型，S-B校正法主要针对的是MLR估计方法的结果。

- M*plus* 在多水平回归模型中，使用MLR、MLM和WLSM估计带有随机效应的模型，不输出Chi-Square Test of Model Fit。此时可以使用S-B校正法输出的校正因子对LL加以校正，然后进行模型比较检验。
- 对于使用MLMV、WLSMV和ULSMV估计的一般模型，差异检验则需要在SAVEDATA和ANALYSIS命令中使用DIFFTEST选项（Muthén & Muthén，2012：451-452）。不过，MLMV不能应用于多水平回归模型。在多水平回归模型中，当ANALYSIS含TYPE＝RANDOM（即含随机斜率）时，不能用WLSMV和ULSMV进行估计。而且，DIFFTEST选项对于TYPE＝TWOLEVEL也不可用。

6.10.1 使用 x^2

S-B调整卡方统计量：

$$\text{Scaled}\,\chi^2 = \frac{(F_0 \times c_0 - F_1 \times c_1)(d_0 - d_1)}{c_0 \times d_0 - c_1 \times d_1} \qquad (6.53)$$

其中，F_0 ＝约束模型 x^2 值；

F_1 ＝自由估计模型 x^2 值；

c_0 ＝约束模型比例因子（scaling factor）；

c_1 ＝自由估计模型比例因子；

d_0 =约束模型自由度；

d_1 =自由估计模型自由度。

注意，在M*plus* 中，对嵌套模型进行卡方检验时需要对模型设定进行调整。以带有随机效应的单因素方差分析和以均值为结果的回归模型进行比较为例，代码（均使用MLR进行估计）见框6-11。

框6-11　对嵌套模型进行卡方检验时模型设定调整的示例

```
MODEL:  ! One-way Anova with Random Effects
  %WITHIN%
  pv1math;
  %BETWEEN%
  pv1math ON sector@0;

MODEL:  ! Means-as-Outcomes Regression
  %WITHIN%
  pv1math;
  %BETWEEN%
  pv1math ON sector;
```

输出结果见框6-12：

框6-12　带有嵌套关系的模型卡方检验输出结果

```
! One-way Anova with Random Effects

Chi-Square Test of Model Fit
Value9.699*


Degrees of Freedom  1
P-Value      0.0018
Scaling Correction Factor        0.8401
  for MLR

! Means-as-Outcomes Regression

Chi-Square Test of Model Fit

Value0.000*


Degrees of Freedom  0
P-Value      0.0000
Scaling Correction Factor        1.0000
  for MLR
```

根据上述结果：

$$F_0 = 9.699，c_0 = 0.8401，d_0 = 1;$$

$F_1 = 0$，$c_1 = 0.9677$，$d_1 = 0$.

代入式6.53计算得：

$$\text{Scaled}\,\chi^2 = \frac{(9.699 \times 0.8401 - 0.000 \times 1.0000)(1 - 0)}{0.8401 \times 1 - 1.0000 \times 0} = 9.699 \qquad (6.54)$$

查卡方分布表，$df = 1$，卡方值9.699，对应p 值为0.00184368。两个模型在0.01水平差异显著。

6.10.2　使用LL计算S-B调整卡方统计量

S-B调整对数似然值统计量：

$$\text{Scaled}\,\chi^2 = -2 \times (L_0 - L_1)\frac{(p_0 - p_1)}{(p_0 \times c_0 - p_1 \times c_1)} \qquad (6.55)$$

其中，L_0 =约束模型对数似然值；

L_1 =自由估计模型对数似然值；

c_0 =约束模型比例因子；

c_1 =自由估计模型比例因子；

p_0 =约束模型参数个数；

p_1 =自由估计模型参数个数。

以随机系数模型和以截距和斜率为结果的回归模型进行比较为例，代码（均使用MLR进行估计）见框6-13。

框6-13　对嵌套模型进行卡方检验时模型设定调整的示例

```
MODEL:          ! Random-Coefficients Regression Model
  %WITHIN%
  s|pv1math ON hisei;
  %BETWEEN%
  pv1math ON sector@0;
  s ON sector@0;
  pv1math with s;

MODEL:        ! Intercepts-and Slopes-as-Outcomes Model
  %WITHIN%
  s|pv1math ON hisei;
  %BETWEEN%
  pv1math ON sector;
  s ON sector;
  pv1math with s;
```

输出结果见框6-14。

框6-14　带有嵌套关系的多水平回归模型输出的LL值

```
! Random-Coefficients Regression

Number of Free Parameters  6

Loglikelihood

H0 Value－27878.507
H0 Scaling Correction Factor    0.9682
for MLR

! Intercepts-and Slopes-as-Outcomes

Number of Free Parameters  8

Loglikelihood

H0  Value－27876.335
H0  Scaling Correction Factor    0.9481
for MLR
```

根据上述结果：

$$L_0=-27878.507，c_0=0.9682，p_0=6;$$

$$L_1=-27876.335，c_1=0.9481，p_1=8.$$

代入式6.55计算得：

$$\text{Scaled}\,\chi^2 = -2 \times [(-27878.507) - (-27876.335)] \times \frac{6-8}{6 \times 0.9682 - 8 \times 0.9481} = 4.8929939 \qquad (6.56)$$

查卡方分布表，$df = 2$，卡方值4.8929939，对应p值为0.08659641。两个模型在0.05水平差异不显著。

一些专门的小软件，如Chi-Square Difference Calculator（CDC5），可以方便地根据M*plus*输出结果进行模型比较检验。也可在R中借助MplusAutomation包来实现模型之间的比较。

6.11　建模策略与一般步骤

6.11.1　模型设定

建模的主要目的是构建结果可解释的简约模型。模型设定需要结合问题实质、研究假设和统计建模多方面的考量。在多水平回归模型中，由于有随机部分，模型设定问题要比一般线性回归模型复杂得多，Snijders＆Bosker（2012）、Raudenbush＆Bryk（2012）等总结了如下指导原则：

- 考虑相关领域知识、既有理论、问题表述和常识。

- 区别研究兴趣关注的、待检验的效应与获得好模型需加入的效应。比如研究关注固定效应，有关随机效应没有既有知识，那么在估计随机部分时，可以采取数据驱动的路线[17]。

● 一般来说，如果模型包括交互效应，那么主效应也应该包括在内（即使不显著）；如果某变量有随机斜率，其固定效应也应该包括在模型之中。

● 是否采取多水平回归模型，可以采取如下策略：（1 ）加入水平-1变量，就要注意组内回归系数可能与组间回归系数不同。可以通过生成一个新变量，等于水平-1变量的组均值，检验该变量的效应。（2）如果某水平-1变量存在较大的随机截距方差，意味着在该变量对结果变量组内效应之间存在较大的未解释的差异。此时，可以在水平-2寻找变量解释这些差异，从而形成跨层交互项。当然，如果理论上有这方面的考虑，也可以加入。

● 注意模型中的随机效应协方差。比如有两个变量有随机效应，那么两个随机斜率之间的协方差也应该包括在模型之内。这样的协方差是模型中的自由参数，需要从数据中进行估计。

● 不必计较模型中包括不显著的变量。它们可能作为控制变量，即使不显著，也可以包括在模型之中。这取决于对研究问题的考虑，或根据研究假设。

● 如果想让模型拟合得更好，只要条件允许，可以纳入增加拟合优度的各个效应。

● 检验各统计参数要控制模型中的其他效应。因为这些效应会影响解释以及统计效力，检验结果可能依赖于它们。统计效力经常较低，意味着检验可能会犯第Ⅰ和第Ⅱ类型错误，尤其是后者。所以一个效应不显著，并不一定意味着在总体中也不存在。反之，一个效应显著，也可能是因为偶然。

● 要在带有合适误差项的模型中检验固定效应，它们可能是随机截距、跨水平交互作用中水平-1变量随机斜率以及水平-1残差项。

● 如果出现增加解释变量R^2减小或者删除解释变量R^2增大的情况，需要考虑模型是否存在误设的问题。

● 除非样本量足够大，水平-1不宜设置很多自变量带有随机斜率，否则迭代难以收敛，且容易产生一大堆不显著的结果。如果有k个随机斜率（加1个随机截距），模型随机部分中就会有（k＋2）（k＋1）/2＋1个参数。

● 一般来讲，很少出现只有随机斜率，而无随机截距的情况；只包括随机斜率而无相应的固定斜率的情况也无意义。

● 水平-1不要包含在本水平不变化的变量。

6.11.2　多水平回归建模的一般步骤

有关多水平回归模型建模的步骤，不同的文献的总结略有差异，但一般遵循如下步骤（Kreft＆Leeuw，2007；王济川等，2008：40）：

①构建空模型，计算组内相关。空模型是多水平建模的基础。只有当数据存在显著的组内相关时，才有必要建立多水平回归模型，否则一般的多元回归即可胜任。因此，多水平建模首先要从构建空模型开始。空模型也构成与其他模型进行比较的基础模型。

②加入水平-2解释变量。如果空模型结果显示存在显著的组内相关，也即意味着存在显著的组间异质性，因此需要在模型中加入组水平变量来解释平均结果的组间变异。如果研究的重点在水平-1中的解释变量，也可以先加入水平-1变量来开始。

③加入水平-1解释变量。加入需分析的水平-1解释变量，或作为背景变量加以控制。一般开始时各解释变量斜率设为固定。

④检验水平-1随机斜率。考虑把水平-1某些变量对应的斜率设为随机，并通过假设检验加以验证。

⑤检验跨水平交互作用。当某些水平-1方程斜率为随机或非随机变动时，即可设定为组群层次变量解释的函数，在斜率方程中加入相应的水平-2变量，在宏观水平上解释其变异。这样在将宏观水平方程带入微观水平方程时，即产生跨水平交互作用项。

6.11.3　构建最终多水平回归模型

前面用PISA数据演示了多水平回归模型的基本原理与各子模型。下面将结合上述建模原则与步骤，结合有关研究和数据情况，构建具有真正分析意义的多水平回归模型。

根据空模型计算ICC的结果表明，就数学成绩问题，存在显著的组内同质性或组间异质性，这表明学生成绩平均结果的组间变异有待解释，使用多水平回归模型分析是必要的。

通过随机系数模型和随机截距模型可以看到，学生数学成绩在各学校之间存在显著差异。水平-2纳入学校类型变量后，一定程度上解

释了结果变量在不同学校之间截距上的变化。如果假定水平-1解释变量与结果变量之间的关系随着不同组群的变化而变化，就可以考虑随机斜率模型。

构建最终模型（final model），需要同时考虑水平-1和水平-2变量，加入能更多解释不同学校在学生数学成绩差异方面的变量，尤其是可能对学生社会经济地位与数学成绩关系起着调解作用的因素。

理论上当然是加入的变量越多解释力度越大，但引入过多变量，尤其是随机效应，会增加模型运行的时间，往往会导致迭代不收敛的问题。组水平解释变量个数也受到不能超过组群个数的限制。

考虑在水平-2加入一些对学生成绩可能产生影响的在学校层面的变量，如学校规模（schsize）、师生比（stratio）、数学师资短缺情况（mtshort，编码自sc11q02）、教学质量（scmatedu）、学生行为（studbeha）、教师参与（tchparti）（可视为教师士气的代理变量）等。

在水平-1，有学生家庭的社会经济地位（组对中处理，cgrouphisei）。同时，水平-2加入作为场景变量的各学校学生家庭平均的社会经济地位（grouphisei）[(18)]，这样结合水平hisei的组中心化值，可分析学生家庭社会经济地位的“情境效应”（contextual effect）。另外，水平-1加入控制变量，如性别（girl）、家庭文化资源（cultposs）等。

进一步可以考虑跨水平交互作用，这尤其需要有一定的研究假设作为指导，比如假定资源多、教学质量好的学校，可以降低学生因家庭社会经济地位差距而在数学成绩上存在的差异，就可以考虑在模型

中纳入这些变量之间的跨层交互项。基于上述考虑，下面构建一个相对完善的“最终模型”：

水平-1：

$$Y_{ij} = \beta_{0j} + \beta_{1j}\text{girl}_{ij} + \beta_{2j}\text{cgrouphisei}_{ij} + \beta_{3j}\text{cultposs}_{ij} + \varepsilon_{ij} \quad (6.57)$$

水平-2：

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{mtshort}_{j} + \gamma_{02}\text{excuract}_{j} + \gamma_{03}\text{tchparti}_{j} + \gamma_{04}\text{cgroupschoolhisei}_{j} + \mu_{0j} \quad (6.58)$$

$$\beta_{1j} = \gamma_{10} + \mu_{1j} \quad (6.59)$$

$$\beta_{2j} = \gamma_{20} + \mu_{2j} \quad (6.60)$$

$$\beta_{3j} = \gamma_{30} + \gamma_{31}\,tchparti_{j} + \mu_{3j} \quad (6.61)$$

整理为混合模型即：

$$\begin{aligned}Y_{ij} = {} & \gamma_{00} + \gamma_{01}\text{mtshort}_j + \gamma_{02}\text{excuract}_j + \\ & \gamma_{03}\text{tchparti}_j + \gamma_{04}\text{cgroupschoolhisei}_j + \gamma_{10}\text{girl}_{ij} + \\ & \gamma_{20}\text{cgrouphisei}_{ij} + \gamma_{30}\text{cultposs}_{ij} + \gamma_{31}\text{cultposs}_{ij} \times \text{tchparti}_j + \\ & \mu_{0j} + \mu_{1j}\text{girl}_{ij} + \mu_{2j}\text{cgrouphise}_{ij} + \mu_{3j}\text{cultposs}_{ij} + \varepsilon_{ij} \qquad (6.62)\end{aligned}$$

使用M*plus* 进行拟合，结果如表6-8所示，具体代码此处省略，可加入本书读者QQ群获取。

表6-8　数学成绩的多水平回归分析

	(1)	(2)	(3)	(4)	(5)
Fixed Effect					
Intercepts (γ_{00})	486.224***	247.985***	500.021***	288.049***	285.903***
	(3.762)	(20.319)	(3.841)	(21.120)	(21.062)
girl (γ_{10})			−18.711***	−19.751***	−19.681***
			(2.137)	(2.141)	(2.203)
cgrouphisei (γ_{20})			0.934***	0.944***	0.941***
			(0.076)	(0.076)	(0.076)
cultposs (γ_{30})			15.908***	15.263***	14.680***
			(1.281)	(1.263)	(1.263)
mtshort (γ_{01})		−8.135*		−7.499*	−7.473*
		(3.416)		(3.488)	(3.483)
excuract (γ_{02})		8.530**		8.322**	8.167**
		(3.033)		(3.084)	(3.083)
tschparti (γ_{03})		5.237**		5.462**	6.211**
		(1.980)		(2.054)	(2.030)
grouphisei (γ_{04})		4.682***		4.168	4.186*
		(0.365)		(0.377)	(0.376)
cultpossXtchparti (γ_{11})					2.736**
					(1.044)
Random Effect					
Intercepts (τ_{00})	1950.856***	652.952***	1842.173***	722.574***	715.052***
	(247.244)	(96.768)	(249.822)	(122.576)	(120.816)
girl (τ_{11})			3.173	12.583	53.195
			(73.130)	(74.747)	(84.107)
cgrouphisei (τ_{22})			0.077	0.077	0.072
			(0.095)	(0.096)	(0.094)
cultpross (τ_{33})			41.708	35.587	26.577
			(26.755)	(25.613)	(24.621)
Residual (σ^2)	5691.353***	5691.624***	5099.222***	5099.291***	5090.483***
	(118.515)	(111.178)	(111.083)	(111.148)	(110.781)
Model Fit					
N	4763	4763	4763	4763	4763
LL	−27536.015	−27465.349	−27293.383	−27234.998	−27231.585
AIC	55078.029	54944.699	54610.766	54501.995	54497.170
BIC	55097.435	54989.979	54688.390	54605.494	54607.137

+ $p<0.1$，* $p<0.05$，** $p<0.01$，*** $p<0.001$

从输出结果来看，根据空模型（1），ICC为1950.856／（1950.856＋5691.353）＝0.25527384，也就是1／4的学生数学成绩方差需要以校际差异来解释，表明有必要使用多水平回归模型进行分析。

加入水平-2变量解释学校间的变异，从模型（2）输出的LL、AIC等指标来看，拟合程度有所改善。简约起见，模型（2）中只保留了显著的变量。模型比较，x^2＝141.330，或者直接使用LL进行比较，D_0-D_1＝－2［（－27536.015）－（－27465.349）］＝141.332，df＝4，差异显著。

模型（3）为随机系数模型。没有纳入水平-2变量。此时，估计水平-1变量方差解释比例，采用R＆B法，与空模型相比，1－5099.222／5691.353＝0.10404046。采取S＆B法，则为1－（1842.173＋5099.222）／（1950.856＋5691.353）＝0.09170307。也就是说，加入性别、家庭社会经济地位、文化资源水平-1变量，使得数学成绩校内方差减少了10％左右。

对于模型（4），若以随机系数模型（3）为基准，纳入水平-2变量后，解释力增加。使用R＆B方法，1－722.574／1842.173＝0.60775997。使用S＆B方法，经前述计算，学生数学成绩各校均值的方差为2132.134；对于模型（4），5099.291／32＋722.574＝881.92684。因此，计算水平-2方差解释比例为1－881.92684／2132.134＝0.58636425。也就是说，纳入模型的若干变量解释学校之间数学成绩差异的比例为60％左右。

水平-2方程中加入学校层面的解释变量后，即可与水平-1方程中的解释变量形成跨水平的交互项。简化起见，只保留显著的交互项。

模型（4）、（5）相比较，$D_0 - D_1 = -2$［（－27234.998）－（－27231.585）］＝6.826，$df = 17 - 16 = 1$，$p = 0.00898402$，在0.01水平统计差异显著。

从结果来看，学生数学成绩的差异，学校因素大概占1／4的比例。学校整体的一些因素对学生数学成绩都有显著影响。学校师资越强、学生家庭平均社会经济地位越高，学生平均数学成绩也越高。可以看到，在学校层面上，学生平均社会经济地位每高1个单位，学生平均数学成绩高4.186分。控制这些因素，从结果来看，学生家庭社会经济地位与数学成绩之间呈现正向关系。相对于所在学校学生平均家庭社会经济地位，平均而言，学生社会经济地位每增加1个单位，学生数学成绩高0.940分。大致而言，多水平框架下，在控制了一些相关变量的情况下，4.186分即代表组间效应，而0.941分即代表组内效应。计算“情境效应”大致为4.186－0.941＝3.245。

本次分析结果表明，学校类型因素（如公立、私立）对于学校的平均数学成绩有显著影响，不过总体上并未支持对学生家庭社会经济地位与数学成绩之间关系有显著调节作用的假设，而教师参与与学生家庭文化资源的跨水平交互作用统计上显著。

6.12　本章小结

实际调查数据大都存在难以忽视的嵌套结构，使用一般回归模型估计就会存在问题。多水平回归模型契合嵌套数据结构，可将宏观、微观因素同时纳入模型，并对跨水平交互作用进行分析。

在简述多水平回归模型的应用背景、基本原理的基础上，本章重点介绍了使用M*plus* 进行多水平回归建模的语法规则、常见子模型、基本步骤、结果解释和相关注意事项。

多水平回归模型是最基本的多水平分析模型。不过在经典的多水平回归模型中，并不包含潜变量概念，所谓情境变量也是由水平-1变量直接聚合而得。这与多水平结构方程模型不同。后者将在后续章节进行介绍。

※推荐阅读※

作为多水平回归模型替代性的方法，稳健标准误回归、固定／随机效应模型在计量经济学中应用较多。固定／随机效应模型主要用于面板数据的分析。可参见伍德里奇（2014，459-481）。尤其注意第三节“把面板数据方法用于其他数据结构”。

作为多水平回归模型入门，可选Kreft＆Leeuw（1998）（中文版2007年出版），其中介绍到整体回归、聚合回归、脉络模型、Cronbach模型等。

多水平回归模型文献较多，经典文献推荐阅读Goldstein（2003）、Raudenbush＆Bryk（2002）（中文版参见Raudenbush＆Bryk，2007）、Gelman＆Hill（2006）；Snijders＆Bosker（2012），以及Hox（2010），其中包括多水平回归模型的一些拓展，如结果变量为二分变量的情况。国内出版的多水平回归模型教材如温福星（2009）、王济川等（2008），也可作为参考。

多水平回归模型应用范围十分广泛。就本章涉及的主题，经典研究如Lee & Bryk（1989）、Gamoran（1992）、Kreft（1997）、Sampson et al（1997）、Ross et al（2000）等，晚近的研究如Chaux et al（2009）、Paulin&Griffin（2016）等。

---

(1)　其他假设还包括：不存在完全共线性、随机抽样、解释变量无测量误差等。

(2)　一般来讲，独立同分布假定规定误差项独立且同分布，但仍然无法确定什么样的分布。正态性假定保证了在小样本数据的情况下，进行统计检验。对于大样本来说，根据中心极限定理，即使误差项不满足正态分布，仍可以对回归参数估计值进行统计推断。另外，在正态分布假定下，最大似然估计与最小二乘法估计结果一致。最大似然估计在大样本情况下具有渐进性质。在样本趋近于无穷大时，不仅满足一致性，而且能够一致估计量中的最小方差。

(3)　所谓“生态谬误”发生在研究者以一个较高层次的或集合形式的分析单位收集资料，但是却想要对一个较低层次的、非集合形式的单位得出论述之时。这个概念最早见于Robinson（1950）的文章。Robinson分析了1930年美国人口普查结果中48个州的识字率与新移民人口比例的关系，发现两者之间相关系数为0.53，即州的新移民比率越高，平均来说这个州的识字率便越高，但当分析个体资料时发现相关系数是－0.11，即平均来说新移民比本地人的识字率低。相反，如果依据个体的特质而推论群体的特质，则犯了“还原谬误”。由于个体层面的两个变量之间的关系可能和群体层面两个变量对应的因素之间的关系不同。比如，针对个体层面的研究发现收入水平和冠心病死亡率之间存在负向的关系，如果由此推论在国家层次，人均国民收入增加与冠心病死亡率呈负相关，就可能犯还原论的错误。因为在不同国家之间，人均国民收入和冠心病死亡率之间可能是正相关关系。

(4)　2009年数据可以从官网下载（本书查阅于2017.1．1）。

(5)　需要注意的是，个体层面变量聚合至组群层面需注意其适当性，需要检验组内一致性和组间异质性的存在。缺乏组内一致性，则个体数据的整合不具有意义；缺乏组间异质性，则无法检验出场景变量的存在（方杰等，2010）。一般来讲，尤其对于参照点位于水平-2的变量，也就是说具有所谓共享性（shared properties）特征的变量（Kozlowski & Klein，2000；Klein & Kozlowski，

2000）或者所谓氛围（climate）变量（Marsh et al，2012），在进行聚合之前，有必要进行组内成员的一致性和均值作为聚合结果的信度的评价。常用指标包括：$R_{wg(j)}$ 、ICC（1）（即ICC）、ICC（2）等（可参见Bliese，2000；Dixon＆Cunningham，2006等概述）。方便起见，相关计算可以在R中使用multilevel包中的ICC1、ICC2、R_{wg} 等命令实现。

(6) 对于式6.6，如果不再假定 α_j 是一系列固定的数，而是有特定概率分布的随机变量，通常 $\alpha_j \sim N(0, \sigma^2)$ 且 $Cov(\alpha_j, X_{ij}) = 0$，另 $\varepsilon_{ij} \sim N(0, \sigma^2)$，可以得到随机效应模型（random effects model）。

(7) 常用的ANOVA与FEM模型较为相似。区别在于FEM模型包含 βX 项模型线性趋势。实际上，如果不指定解释变量，二者几乎就是一样的。

(8) 当然，多水平回归模型若要取得优良的估计也必须满足若干假定，可参见相关文献介绍（如Raudenbush＆Bryk，2002：241）。

(9) 本质上，多水平回归模型的优势之所以成立，在于多水平回归模型构建了适应嵌套数据结构的误差结构，考虑到不同水平的变异。与单水平模型相比，区分了不同层次误差的来源，将单一的随机误差项分解到与数据层次结构相应的各水平上，更好地分解了方差-协方差成分，并提供了有效估计（廖卉等，2012）。

(10) 在多水平结构方程模型中，这种处理其实更属于一般情况。同时出现WITHIN和BETWEEN不同部分的同一变量名所指代的，实际上对应的是该变量分解后的组内和组间不同（潜在）部分（详见第8章 ）。

(11) 这是因为，作为结果变量，pv1math没有指定隶属的层次，所以可以在两个水平参与子模型构建。在组间层次，即随机截距。实际上，pv1math在水平-1的方程截距的随机变化，表现在水平-2，在组间不同，方差存在于水平-2。按照多水平结构方程模型的框架，同时出现WITHIN和BETWEEN不同部分的同一变量名所指代的，对应的是该变量分解后的组内和组间不同（潜在）部分。在多水平回归模型中，协变量都需要指定所属层面，而不会出现两用变量。只有结果变量才会“两用”，无需指明所属层面。协变量是否可（在不同水平）两用，正是多水平回归模型和多水平结构方程模型的本质差异。参见第8章 内容。

(12) 中心化可以用于连续变量，也可以用于虚拟变量（Raudenbush＆Bryk，2002：34）。比如，对于性别的虚拟变量（男＝1；女＝0），均值即代表样本中男性的比例。原虚拟变量的0或1值转化为0或1减去样本中男性的比例。当个体为男性时，$X_{ij}-\bar{X}_{..}$ 等于样本中女性的比例；当个体为女性时，$X_{ij}-\bar{X}_{..}$ 等于样本

中男性比例的负数。截距可解释为对应样本中女性比例的结果变量的均值（$X_{ij}=0$时的预测值）。如果按组均值中心化，对于男性，$X_{ij}-\bar{X}_{.j}$ 为第j 个学校中女性的比例；对于女性，$X_{ij}-\bar{X}_{.j}$ 为第j 个学校中男性比例的负数。截距代表第j 个学校的结果变量的均值。

(13)　一些软件（如MLwin）提供了迭代广义最小二乘估计（Iterative Generalized Least Squares，IGLS）和限制迭代广义最小二乘估计（Restricted Iterative Generalized Least Squares，RIGLS），分别对应FIML和REML（Goldstein，2003）算法。二者之间最主要的区别是IGLS使用最大似然值，而RIGLS使用受限最大似然值。

(14)　一般来讲，REML、RIGLS、MCMC等方法，都可以获得无偏估计，GLS与ML法具有相同的效能，得到的γ估计值都是无偏的，不一定用EB这类复杂度高的技术（Kreft＆De Leeuw，1998）。

(15)　需要注意的是，不同的模型设定、不同的因变量类型，需要考虑相应的估计方法。有关多水平回归模型的估计方法参见M*plus* 手册（Muthén＆Muthén，2012：601-602）。

(16)　通常而言，在模型复杂程度和数据状况一定的情况下，模型估计迭代迅速收敛表明该模型拟合数据较好；反之，模型迭代不能收敛则表明该模型设定可能有问题（王济川等，2008：32）。

(17)　建模的过程是理论假设和统计技术相互结合的产物。一般来讲，如果研究有了比较明确的假设，要对纳入模型的变量有所选择和考虑。但实际建模过程中，仍可以采取步进的策略，进行实验性的建模。至少要以空模型作为起点，并通过模型比较检验，突出某些重点解释变量的作用。随机效应一般也是在变量固定部分显著之后考虑加入的。

(18)　模型中纳入学校家长平均社会地位的变量，一定程度上也有利于降低来自不同社会经济地位背景的学生进入不同的学校而造成的样本选择导致的估计偏差（selection bias）。可以尝试的是，如果水平-2不加入各学校学生平均社会经济地位的场景变量，水平-2其他变量回归系数取值要大得多，学生行为（studbeha）也显著。这部分原因在于学生上什么样的学校，是一个有着高度选择性的现象。一般而言，家庭社会经济地位较高的学生更可能上较好的学校，其成绩也相对较高。因此，水平-2加入家庭平均社会经济地位，虽不能完全克服，但一定程度上有助于减低这种选择性导致的估计偏差，否则可能会严重高估学校特征对学生成绩的影响。

7　多水平增长模型

7.1　历时数据及分析框架

7.1.1　历时数据的概念

本章继续使用中国健康与营养调查（CHNS）数据作为示例。这是一个典型的历时（longitudinal）数据，适合增长分析方法。

所谓“历时数据分析”，是研究变迁问题常用的方式，在因果推论方面具有截面数据不具有的优势。历时数据分析具备以下几个方面的特征（Singer＆Willett，2003：9-15）：

● 3个或更多波次（waves）数据。两轮数据无法描述个体发展的轨迹，两个状态之间的静态，不能区别真实的变化和测量误差。由于历时数据针对每个对象反复收集数据，因此研究对象内的观测值之间会存在相关。

● 对时间敏感的度量。选择最能反映结果变量变化规律的时间度量。比如对于学校研究，可以选用年级或年龄。该指标并不一定是时间，比如研究轮胎损耗时，可以选用汽车行驶里程数。但需要满足单向性。比如对于儿童的研究，可以用身高而不能用体重，作为时间的量度。

■ 数据收集波次的间隔可以相等（equally spaced waves），形成平衡的数据。但有时为了更好地刻画个体发展轨迹，也可以不等。

■每个个体收集数据可以按照同样的时序表，形成所谓“固定时间结构”（time-structured）数据；也可以在不同个体之间变化，形成“非时间结构”（time-unstructured）数据。

■各个体测量波次可以一致，即所谓均衡（balanced）数据。然而实际之中，由于失访、参与者损耗等原因，测量波次也可以不一致，形成所谓非均衡（unbalanced）或不完整（incomplete）数据。在随机缺失（missing at random）的情况下，增长具有处理非平衡数据和不完整数据的能力。可在最大似然或限制性最大似然的基础上，利用全部可以利用的数据进行模型估计。无需剔除带有缺失值的研究对象，也不需要对缺失值进行插补。

● 结果变量的取值随着时间系统性地变化。历时研究的测量必须保证有效性和准确性，同时还要满足跨时间等值性，在历次测量中具有相等的效度和精确性。

历时数据研究面临着有关变迁的两种类型问题：

①结果变量是如何随着时间变化的？该问题是描述性的，需要我们刻画每个个体随着时间变化的模式。个体变化是线性的还是非线性的？变化模式是稳定的还是有波动的？这对应着水平-1的分析，即个体内部随时间的变化（within-individual change），描述每个个体自身成长轨迹的形状。

②我们能预测个体之间这些变化的差异吗？该问题是解释性的，需要我们检验预测变量与变化模式的关系。哪个解释变量与哪个模式是关联的？这对应着水平-2的分析，即个体之间变化的差异

（interindividual differences in change），需要分析不同个体变化的异质性，并解释预测变量与个体发展轨迹之间的关系（Singer＆Willett，2003：9-15）。

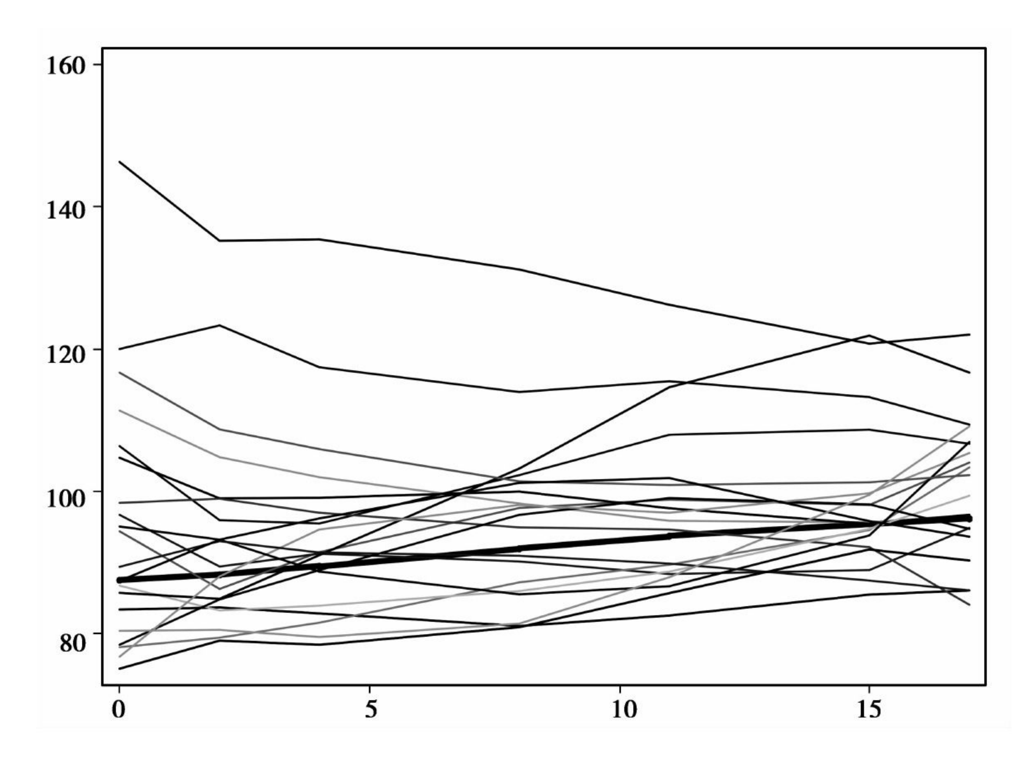


图7-1　采用非参数方法平滑化的个体血压变化轨迹差异

（调查波次7次的前20位个体）

7.1.2　基于多水平框架的增长模型

本章介绍的是基于多水平框架的增长模型，也就是将不同次测量视作嵌套在个体内部，从而可以使用多水平回归模型进行建模。

水平-1：

$$Y_{ti}=\pi_{0i}+\pi_{1i}T_{ti}+\pi_{2i}X_{ti}+e_{ti} \quad (7.1)$$

水平-2：

$$\pi_{0i}=\beta_{00}+\beta_{01}W_i+r_{0i} \quad (7.2)$$

$$\pi_{1i}=\beta_{10}+\beta_{11}W_i+r_{1i} \quad (7.3)$$

$$\pi_{2i}=\beta_{20}+\beta_{21}W_i+r_{2i} \quad (7.4)$$

其中，T_{ti} 是观测时间的指示变量，如［0，1，2，3］；每个个体的初始水平 π_{0i} 和增长轨迹 π_{1i} 分别以期望均值 β_{00} 、 β_{10} ，随机偏差 r_{0i} 、 r_{1i} 的线性组合表示。β_{00} 、 β_{11} 分别表示非时变变量 W_i 对于初始水平和增长轨迹的效应， π_{2i} 代表了时变协变量（time-varying covariates）X_{ti} 的效应。

将方程7.2、7.3、7.4带入方程7.1，得复合模型：

$$Y_{ti}=\beta_{00}+\beta_{01}W_i+\beta_{10}T_{ti}+\beta_{20}X_{ti}+\beta_{11}W_i\times T_{ti}+\beta_{21}W_i\times X_{ti}+r_{0i}+r_{1i}T_{ti}+r_{2i}X_{ti}+e_{ti} \quad (7.5)$$

研究变迁问题，除了采取多水平回归模型，也可以采取结构方程模型或协方差结构分析的框架，构成所谓“潜增长模型”（latent growth model），见本书第4章 、第5章 。

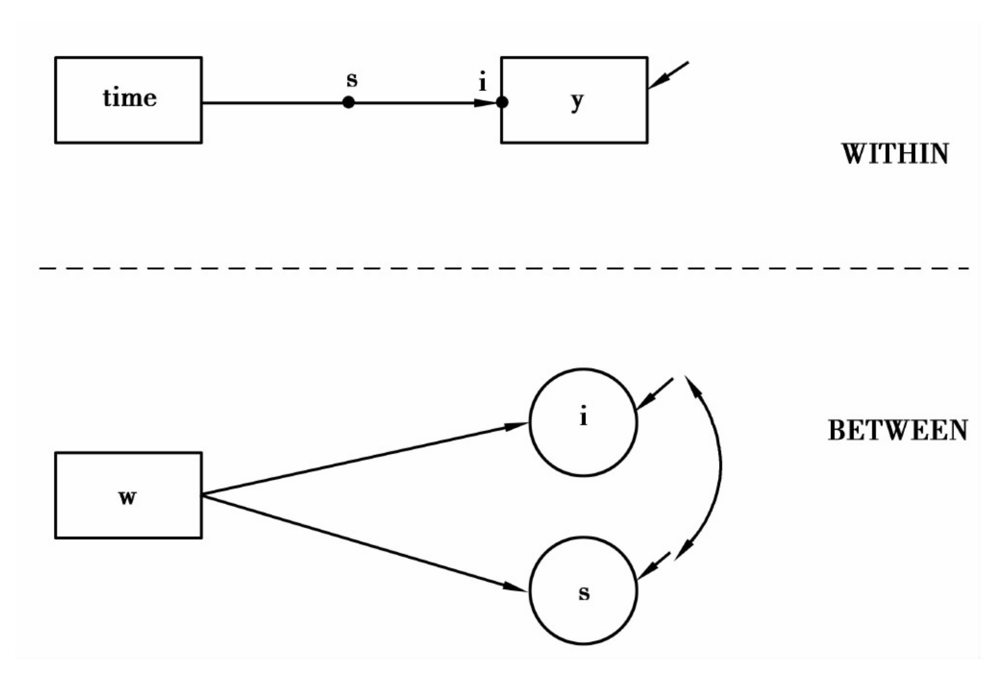


图7-2 多水平增长模型

多水平增长模型中，时间作为一个最低水平的自变量，个体在水平-2定义，解释变量可以在各层纳入模型。刻画增长曲线的参数上个体之间的差异作为随机效应而被模型化。结构方程模型中，采取潜变量视角，时间外在地整合进潜变量的设定之中。个体曲线的参数即初始水平和线性变化速率作为潜变量，带有协方差和均值结构。

测量方程部分：

$$Y_{ti} = \lambda_{0t}\ \eta_{0i} + \lambda_{1t}\ \eta_{1i} + \gamma_{2t} X_{ti} + \varepsilon_{ti} \qquad (7.6)$$

结构方程部分：

$$\eta_{0i} = \nu_{0} + \gamma_{0} W_i + \zeta_{0i} \qquad (7.7)$$

$$\eta_{1i} = \nu_{1} + \gamma_{1} W_i + \zeta_{2i} \qquad (7.8)$$

其中，Y_{ti} 代表个体i在t次测量中的结果；初始水平和成长轨迹由潜变量η_{0i}和η_{1i}表示，各自期望均值为ν_0和ν_1，随机部分为ζ_{0i}和ζ_{2i}；时间的因素通过对因子载荷λ_{0t}和λ_{1t}进行约束得到，分别为［1，1，1，1］和［0，1，2，3］；γ_{2t}代表了时变协变量X_{ti}的效应；γ_0和γ_1分别代表了非时变协变量对于初始水平和增长轨迹的效应。

在多水平回归模型中，时变协变量可有随机效应系数，系数在不同个体之间变化；在SEM潜增长模型中，时变协变量有固定效应系数，系数随时间而变化（Bliese，2010）。在多水平回归模型中，假设每次测量有相同的测量误差。而在潜增长曲线模型中，可对每次测量误差分别估计。在处理变量间复杂关系时后者也更为灵活。

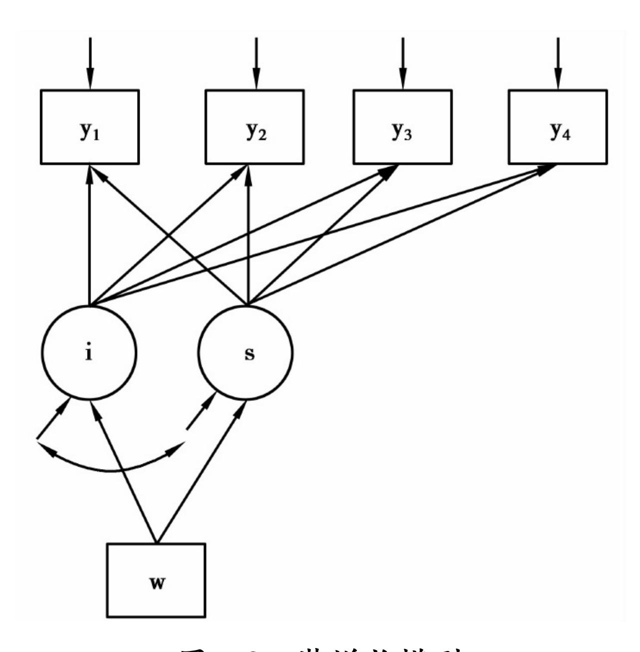


图7-3　潜增长模型

实际上，潜增长模型中的潜变量对应于多水平增长模型中的随机效应。如果能设定同样的模型，两种模型将得到基本一致的参数估计

结果（Stoel et al，2003）。二者数学上是等同的，差别主要是统计软件和历史传统造成的。随着统计软件的发展，彼此的局限性正在消失。甚至在M*plus* 中，可将多水平模型和潜增长模型结合起来使用（参见Muthén＆Asparouhov，2011）。从实际经验来看，有些潜增长模型建模运算有问题的模型，使用多水平回归模型建模可能会得到结果。

7.1.3 数据结构

在M*plus* 中，增长模型可以允许以长型（LONG）或宽型（WIDE）两种方式建模，分别对应潜增长模型（LGM）和多水平增长模型（MLM）。本节所用CHNS血压数据的长型格式见框7-1。

框7-1 长型数据

id	wave	time	age	press	gen~r	bmi
1	89	0	43	106.7	1	21.1
1	91	2	45	101.1	1	23.2
1	93	4	47	85.0	1	22.6
2	89	0	42	86.7	0	19.9
2	91	2	44	90.0	0	20.1
2	93	4	46	84.4	0	21.2
3	89	0	40	100.0	1	24.3
3	91	2	42	97.2	1	.
3	93	4	43	86.1	1	25.2
4	00	11	52	144.4	1	23.5
4	04	15	56	106.4	1	24.9
4	06	17	58	100.9	1	25.4
4	89	0	41	90.0	1	22.3
4	91	2	43	96.7	1	23.8
4	93	4	45	95.0	1	23.4
5	00	11	50	90.2	1	20.6
5	04	15	55	100.7	1	21.5
5	06	17	56	84.7	1	21.1
5	89	0	40	83.3	1	17.8
5	91	2	42	83.3	1	19.8
5	93	4	44	94.4	1	18.2

对应的宽型格式见框7-2。

框7-2　宽型数据（分两行显示）

```
+-----------------------------------------------------------------
   id    pr~00   pr~04   pr~06   pr~89   pr~91   pr~93   pr~97
 -----------------------------------------------------------------
    1       .       .       .    106.7   101.1    85.0       .
    2       .       .       .     86.7    90.0    84.4       .
    3       .       .       .    100.0    97.2    86.1       .
    4   144.4   106.4   100.9     90.0    96.7    95.0       .
    5    90.2   100.7    84.7     83.3    83.3    94.4       .
+-----------------------------------------------------------------
--------------------------------------------------------------------+
 bmi00   bmi04   bmi06   bmi89   bmi91   bmi93   bmi97   gen~r
--------------------------------------------------------------------
     .       .       .    21.1    23.2    22.6       .       1
     .       .       .    19.9    20.1    21.2       .       0
     .       .       .    24.3       .    25.2       .       1
  23.5    24.9    25.4    22.3    23.8    23.4       .       1
  20.6    21.5    21.1    17.8    19.8    18.2       .       1
--------------------------------------------------------------------+
```

统计软件进行多水平增长模型的分析时，需要输入长型数据，然而数据录入时多为宽型数据，所以需要进行格式转换。很多具有数据管理功能的统计软件提供了长型数据与宽型数据之间互相转换（restructure）的命令，如Stata中的reshape、R中的reshape2、SPSS中的restructure等。M*plus* 中，数据之间转换的命令为：DATA WIDETOLONG和DATA LONGTOWIDE。

M*plus* 数据重构命令见框7-3。

框7-3 M*plus* 数据重构命令

```
DATA WIDETOLONG:    ! 宽型数据转换为长型
WIDE=names of old wide format variables;
LONG=names of new long format variables;
```

```
IDVARIABLE＝name of variable with ID information;
REPETITION＝name of variable with repetition information;

DATA LONGTOWIDE:      ！长型数据转换为宽型
LONG＝names of old long format variables;
WIDE＝names of new wide format variables;
IDVARIABLE＝name of variable with ID information;
REPETITION＝name of variable with repetition information;
```

以CHNS血压数据为例，采用上述命令，对CHNS血压数据进行由宽型变成长型的格式转换，数据重构命令见框7-4。

框7-4　M*plus* 数据重构命令

```
TITLE: WIDE TO LONG;
Data: File is chnspress_wide.dat;
DATA WIDETOLONG:
   WIDE ＝ press89  press91  press93  press97  press00  press04
press06|
     bmi89 bmi91 bmi93 bmi97 bmi00 bmi04 bmi06;
  LONG＝press|bmi;
  IDVARIABLE＝id;
  REPETITION＝wave;
```

7.2　线性增长模型

在多水平框架下，增长模型水平-1子模型是测量水平，描述个体随时间的发展：

$$Y_{ti} = \pi_{0i} + \pi_{1i}a_{ti} + \pi_{2i}a_{ti}^2 + \cdots + \pi_{pi}a_{ti}^p + e_{ti} \tag{7.9}$$

其中，Y_{ti} 为个体水平结果变量；i ＝1，2，…，n 代表不同的个体；a_{ti} 指的是个体在时间 t 上的测度；π_{pi} 则对应于个体增长曲线的参数 p ，与多项式的阶数相联系（p ＝0，1，…，P ）；e_{ti} 为误差项，可以有简单或复杂的结构。

水平－2子模型，个体水平，解释个体间增长差异：

$$\pi_{pi} = \beta_{p0} + \sum_{q=1}^{Qp} \beta_{pq}X_{qi} + r_{pi} \tag{7.10}$$

其中，X_{qi} 是个体背景变量，或实验分组；β_{pq} 表示 X_q 对第 p 个增长参数的效应；r_{pi} 为平均值为零的随机效应（Raudenbush＆Bryk，2002：162）。

7.2.1　无条件均值模型

最简单的线性增长模型即无条件均值模型（unconditional means model），相当于一般多水平回归模型中的空模型。

水平-1子模型：

$$Press_{ti} = \pi_{0i} + e_{ti} \quad (7.11)$$

水平-2子模型：

$$\pi_{0i} = \beta_{00} + r_{0i} \quad (7.12)$$

复合模型：

$$Press_{ti} = \beta_{00} + (r_{0i} + e_{ti}) \quad (7.13)$$

其中，$e_{ti} \sim N(0, \sigma^2)$，$r_{0i} \sim N(0, \tau_{00})$。相应M*plus* 代码见框7-5。

框7-5　无条件均值模型

```
TITLE: Unconditional Means Model;
Data:
  File is chnspress.dat;
Variable:
  Names are id time age press gender bmi wave;
  USEVARIABLES＝id press;
  Missing are all (－9999) ;
  WITHIN＝;
  BETWEEN＝;
  CLUSTER＝id;
ANALYSIS: TYPE＝TWOLEVEL;
          ESTIMATOR＝ML;
```

```
MODEL:
  %WITHIN%
     press;
  %BETWEEN%
     press;
```

运行上述代码，输出结果见框7-6。

框7-6　无条件均值模型结果（部分）

```
Estimated Intraclass Correlations for the Y Variables
                   Intraclass
     Variable      Correlation
     PRESS            0.370
 MODEL RESULTS
                                                       Two-Tailed
                    Estimate      S.E.      Est./S.E.     P-Value
 Within Level
 Variances
     PRESS           106.931      2.496       42.840        0.000
 Between Level
 Means
     PRESS            91.126      0.316      288.129        0.000
 Variances
     PRESS            62.961      4.126       15.259        0.000
```

无条件增长模型作为基准模型，将结果变量总方差拆分为个体内方差（with-person）σ^2 和个体间方差（between-person）τ_{00}，由此可以计算ICC。对于增长模型而言，ICC指的是结果变量总的方差当中个体间方差所占比例。*Mplus* 无条件增长模型结果已经输出了ICC，当然也可以自己计算：

$$ICC = \frac{\tau_{00}}{\tau_{00} + \sigma^2} = \frac{62.961}{62.961 + 106.931} = 0.37059426 \qquad (7.14)$$

无条件均值模型下的ICC也展现了残差自相关的大小。从结果来看，本例中该队列个体间方差所占比例个体各时期血压变化具有加高的组内相关。

7.2.2　无条件增长模型

在无条件均值基础上，水平-1方程纳入时间维度，即构成无条件增长模型（unconditional growth model）。

水平-1子模型：

$$Press_{ti} = \pi_{0i} + \pi_{1i} Time_{ti} + e_{ti} \qquad (7.15)$$

水平-2子模型：

$$\pi_{0i} = \beta_{00} + r_{0i} \qquad (7.16)$$

$$\pi_{1i} = \beta_{10} + r_{1i} \qquad (7.17)$$

复合模型：

$$Press_{ti} = \beta_{00} + \beta_{10} Time_{ti} + (r_{0i} + r_{1i} Time_{ti} + e_{ti}) \qquad (7.18)$$

其中，$e_{tj} \sim N(0, \sigma^2)$，

$$\begin{pmatrix} r_{0i} \\ r_{1i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix} \right)$$

在M*plus* 中，该模型对应的代码见框7-7。

框7-7　无条件增长模型

```
TITLE: Unconditional Growth Model;
Data:
  File is chnspress.dat;
Variable:
  Names are id time age press gender bmi wave;
  USEVARIABLES＝id press time;
  Missing are all（－9999）;
  WITHIN＝time;
  BETWEEN＝;
  CLUSTER＝id;
ANALYSIS: TYPE＝TWOLEVEL RANDOM;
          ESTIMATOR＝ML;
MODEL:
  %WITHIN%
    s | press ON time;
  %BETWEEN%
    press;
```

```
s;
press with s;
```

运行上述代码，输出结果见框7-8。

框7-8　无条件增长模型输出结果（部分）

```
MODEL RESULTS
                                                        Two-Tailed
                     Estimate     S.E.     Est./S.E.     P-Value
 Within Level
Residual Variances
    PRESS              81.811     2.114       38.696       0.000
Between Level
 PRESS                 WITH
     S                 -1.328     0.371       -3.582       0.000
 Means
     PRESS             87.489     0.356      245.674       0.000
     S                  0.551     0.032       17.251       0.000
 Variances
     PRESS             70.318     5.334       13.183       0.000
     S                  0.315     0.039        8.110       0.000
```

通过比较无条件均值模型和无条件成长模型，可以呈现加入时间变量后个体内部被时间因素解释的变差比例（Singer＆Willett，2003：100）。

$$\text{Pseudo}\quad R_t^2 = \frac{\hat{\sigma}^2_{\text{无条件增长模型}} - \hat{\sigma}^2_{\text{无条件均值类型}}}{\hat{\sigma}^2_{\text{无条件均值类型}}} = \frac{106.931 - 81.811}{106.931} = 0.23491784 \tag{7.19}$$

从结果来看，个体血压的线性变化中23.5％可由年龄的增加来解释。另外，无条件成长模型中，水平-2残差协方差反映了（未控制其

他协变量的情况下）初始状态和变化速率之间的关系。[1]比如，对于个体血压变化的问题，该指标意味着40—50岁队列开始血压较高者今后血压高增加得快（或者慢）。

$$\hat{\rho}_{\pi_0\pi_1} = \hat{\rho}_{01} = \frac{\hat{\tau}_{01}}{\sqrt{\hat{\tau}_{01}\ \hat{\tau}_{11}}} = \frac{-1.328}{\sqrt{70.318 \times 0.315}} = -0.28216916 \qquad (7.20)$$

从结果来看，总体来说，40—50岁初始血压和血压增长速率之间呈负向关系，且统计上显著。

7.2.3　增长模型中的协变量

增长模型中，协变量可以分为两种类型：①非时变协变量（time-invariant covariates），即个体之间不同，但不随时间变化，解释了成长因素的变异；②时变协变量（time-invariant covariates）。个体之间不同，也随着时间变化，解释了超出成长因素之外结果的变异。

纳入非时变协变量

作为示例，下面在模型中纳入性别变量。由于性别是既定的，所以属于非时变协变量。

水平-1子模型：

$$\text{Press}_{ti} = \pi_{0i} + \pi_{1i}\,\text{Time}_{ti} + e_{ti} \qquad (7.21)$$

水平-2子模型：

$$\pi_{0i} = \beta_{00} + \beta_{01} \text{Gender}_i + r_{0i} \quad (7.22)$$

$$\pi_{1i} = \beta_{10} + \beta_{11} \text{Gender}_i + r_{1i} \quad (7.23)$$

复合模型：

$$\text{Press}_{ti} = \beta_{00} + \beta_{01} \text{Gender}_i + \beta_{10} \text{Time}_{ti} + \beta_{11} \text{Gender}_i \times \text{Time}_{ti} + r_{0i} + r_{1i} \text{Time}_{ti} + e_{ti} \quad (7.24)$$

其中，$e_{tj} \sim N(0, \sigma^2)$。

$$\begin{pmatrix} r_{0i} \\ r_{1i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix} \right)$$

M*plus* 代码见框7-9。

框7-9　纳入非时变变量

```
TITLE: Add Time-invariant Covariates;
Data:
  File is chnspress.dat;
Variable:
  Names are id time age press gender bmi wave;
  USEVARIABLES=id press gender time;
```

```
  Missing are all (—9999) ;
  WITHIN=time;
  BETWEEN=gender;
  CLUSTER=id;
ANALYSIS: TYPE=TWOLEVEL RANDOM;
          ESTIMATOR=ML;
MODEL:
  %WITHIN%
  s | press ON time;
  %BETWEEN%
  s press ON gender;
  s with press;
```

运行代码，输出结果见框7-10。

框7-10　纳入非时变变量输出结果（部分）

```
Within Level
 Residual Variances
    PRESS              81.799     2.114       38.699      0.000
 Between Level
 S         ON
    GENDER             -0.059     0.064       -0.918      0.358
 PRESS     ON
    GENDER              1.700     0.710        2.395      0.017
 S         WITH
    PRESS              -1.306     0.369       -3.537      0.000
Intercepts
    PRESS              86.645     0.500      173.295      0.000
    S                   0.580     0.045       12.972      0.000
 Residual Variances
    PRESS              69.609     5.299       13.136      0.000
    S                   0.315     0.039        8.111      0.000
```

对于线性增长模型，在未控制其他变量的情况下，从结果来看，性别因素对于血压的初始状态的解释统计上显著，但并未显著影响血压变化的速率。平均来讲队列女性初始血压为86.645 mmHg，男性和女性初始血压差异为1.700 mmHg。女性血压增长速率为0.580 mmHg，男性和女性血压增长速率差异为－0.059 mmHg，但统计上不显著（p ＝0.358）。

与无条件成长模型相比，个体内方差成分（$\hat{\sigma}^2$）变化不大且依然显著，表明需要加入时变变量对血压变化加以解释。水平-2方差成分τ_{00}、τ_{11}只是略有变化，且显著，暗示仍需要探寻血压初始水平和成长速率方面残差变异的潜在解释。

纳入时变协变量

考虑到血压与BMI指数有关，纳入BMI指数。与性别不同，由于每次测量时人们的身高、体重均会发生变化，所以BMI指数属于时变协变量。

水平-1子模型：

$$\text{Press}_{ti} = \pi_{0i} + \pi_{1i}\text{Time}_{ti} + \pi_{2i}\text{BMI}_{ti} + e_{ti} \quad (7.25)$$

水平-2子模型：

$$\pi_{0i} = \beta_{00} + \beta_{01}\text{Gender}_{i} + r_{0i} \quad (7.26)$$

$$\pi_{1i} = \beta_{10} + \beta_{11}\,\text{Gender}_i + r_{1i} \qquad (7.27)$$

$$\pi_{2i} = \beta_{20} + \beta_{21}\,\text{Gender}_i + r_{2i} \qquad (7.28)$$

复合模型：

$$\text{Press}_{ti} = \beta_{00} + \beta_{01}\,\text{Gender}_i + \beta_{10}\,\text{Time}_{ti} + \beta_{20}\,\text{BMI}_{ti} + \beta_{11}\,\text{Gender}_i \times \text{Time}_{ti} + \beta_{21}\,\text{Gender}_i \times \text{BMI}_{ti} + r_{0i} + r_{1i}\,\text{Time}_{ti} + r_{2i}\,\text{BMI}_{ti} + e_{ti} \qquad (7.29)$$

其中，

$$e_{tj} \sim N(0, \sigma^2)$$

$$\begin{pmatrix} r_{0i} \\ r_{1i} \\ r_{2i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \tau_{00} & \tau_{01} & \tau_{02} \\ \tau_{10} & \tau_{11} & \tau_{12} \\ \tau_{20} & \tau_{21} & \tau_{22} \end{pmatrix} \right)$$

在多水平框架中，个体特定（person-sepecific）预测变量的效应对不同个体而言应是常量。对增长模型而言，非时变变量应符合这个规则。然而，对于时变变量，可以根据研究假设和数据情况，将时变变量设置为不同的情况：

$$\pi_{2i} = \beta_{20} \qquad (7.30)$$

意味着时变变量BMI只是在个体内部随时间而变化。

$$\pi_{2i} = \beta_{20} + r_{2i} \qquad (7.31)$$

意味着时变变量BMI在队列的个体之间随机变化。

$$\pi_{2i} = \beta_{20} + \beta_{21}\text{Gender}_i + r_{2i} \quad (7.32)$$

意味着BMI在个体之间的变化受到非时变变量Gender的影响。

上述模型对应的M*plus* 代码见框7-11。

框7-11　纳入时变变量

```
TITLE: Add Time-varying Covariates;
Data:
  File is chnspress.dat;
Variable:
  Names are id time age press gender bmi wave;
  USEVARIABLES＝id press gender time bmi;
  Missing are all（－9999）;
  WITHIN＝time bmi;
  BETWEEN＝gender;
  CLUSTER＝id;
ANALYSIS: TYPE＝TWOLEVEL RANDOM;
          ESTIMATOR＝ML;
MODEL:
    %WITHIN%
    s1|press ON time;
    s2|press ON bmi;
    %BETWEEN%
```

```
s1  s2 press ON gender;
s1  with press;
s2  with press;
s1  with s2;
```

运行上述代码，输出结果见框7-12。

框7-12　纳入时变变量输出结果（部分）

```
MODEL RESULTS
                                                              Two-Tailed
                      Estimate      S.E.      Est./S.E.        P-Value
Within Level
 Residual Variances
  PRESS                 80.490      2.137        37.662          0.000
Between Level
 S1         ON
  GENDER                -0.050      0.065        -0.772          0.440
 S2         ON
  GENDER                 0.117      0.170         0.690          0.490
 PRESS      ON
  GENDER                 0.109      3.698         0.029          0.977
 S1        WITH
  PRESS                  1.866      1.662         1.123          0.262
  S2                    -0.140      0.078        -1.791          0.073
 S2        WITH
  PRESS                 -5.583      5.484        -1.018          0.309
 Intercepts
  PRESS                 61.562      2.507        24.556          0.000
  S1                     0.493      0.046        10.745          0.000
  S2                     1.108      0.113         9.786          0.000
 Residual Variances
  PRESS                120.229    120.717         0.996          0.319
  S1                     0.317      0.042         7.570          0.000
  S2                     0.362      0.251         1.445          0.148
```

另外，如果假定时变变量BMI也影响增长轨迹的斜率，可以在水平-1模型中增加$Time_{ti}$ 和BMI_{ti} 的交互项。该交互项同样也是时变变

量，不同于混合模型中$Gende_{ri}$ 与$Time_{ti}$ 之间的交互项。如果该项统计上显著，意味着BMI对于血压的影响也会随着年龄的增加而变化。相应代码见框7-13。

框7-13　纳入时变变量与时间的交互项

```
TITLE: Add Time-varying Covariates, with interactions;
Data:   File is chnspress.dat;
Variable:
  Names are id time age press gender bmi wave;
  USEVARIABLES=id press gender bmi time timeXbmi;
  Missing are all (-9999) ;
  WITHIN=time bmi timeXbmi;
  BETWEEN=gender;
  CLUSTER=id;
DEFINE: timeXbmi=time*


bmi;
ANALYSIS: TYPE=TWOLEVEL RANDOM;
          ESTIMATOR=ML;
MODEL: %WITHIN%
     s1|press ON time;
     s2|press ON bmi;
     s3|press ON timeXbmi;
     %BETWEEN%
     s1  s2 press ON gender;
```

```
s3;
s1  with press;
s2  with press;
s3  with press;
s1  with s2;
s2  with s3;
s1  with s3;
```

运行上述代码，输出结果见框7-14。

框7-14　纳入时变变量与时间的交互项输出结果（部分）

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Residual Variances				
PRESS	80.260	2.144	37.436	0.000
Between Level				
S1 ON				
GENDER	-0.066	0.065	-1.007	0.314
S2 ON				
GENDER	0.085	0.169	0.502	0.616
PRESS ON				
GENDER	0.900	3.673	0.245	0.806
S1 WITH				
PRESS	-5.687	23.649	-0.240	0.810
S2	0.585	1.087	0.538	0.590
S3	-0.074	0.113	-0.655	0.513
S2 WITH				
PRESS	-6.781	12.006	-0.565	0.572
S3	-0.030	0.047	-0.623	0.533
S3 WITH				
PRESS	0.295	1.042	0.283	0.777

Means				
S3	-0.021	0.011	-1.998	0.046
Intercepts				
PRESS	58.347	3.013	19.366	0.000
S1	0.985	0.244	4.036	0.000
S2	1.250	0.137	9.153	0.000
Variances				
S3	0.004	0.005	0.738	0.460
Residual Variances				
PRESS	107.307	263.579	0.407	0.684
S1	1.747	2.576	0.678	0.498
S2	0.493	0.545	0.906	0.365

整理上述模型结果到回归表，见表7-1。

表7-1 CHNS1989年40—50岁队列血压线性增长模型

	(1)	(2)	(3)	(4)	(5)
Fixed Effect					
Initial status, π_{0i}					
Intercepts(β_{00})	91.126***	87.489***	86.645***	61.562***	57.347***
	(0.316)	(0.412)	(0.500)	(2.507)	(3.013)
gender(β_{01})			1.700*	0.109	0.900
			(0.815)	(3.698)	(3.673)
Rate of change for $Time_{ti}$, π_{1i}					
Intercepts(β_{10})		0.551***	0.580***	0.493***	0.985***
		(0.065)	(0.045)	(0.046)	(0.244)
gender(β_{11})			−0.059	−0.050	−0.066
			(0.064)	(0.065)	(0.065)
BMIti, π_{2i}					
Intercepts(β_{20})				1.108***	1.250***
				(0.064)	(0.137)
gender(β_{21})				0.117	0.085
				(0.170)	(0.169)
$Timec_{ti}$ X BMI_{ti}, π_{3i}					
Intercepts(β_{30})					−0.021*
					(0.011)
Variance Components					
Within-person(σ^2)	106.931***	81.811***	81.799***	80.490***	80.260***
	(2.496)	(2.114)	(2.114)	(2.137)	(2.144)
Initial status(τ_{00})	62.961***	70.318***	69.609***	120.229	107.307
	(4.126)	(5.334)	(5.299)	(120.717)	(263.579)
Rate of change(τ_{11})		0.315***	0.315***	0.317***	1.747
		(0.039)	(0.039)	(0.042)	(2.576)
续表					
BMI_{ti}(τ_{22})				0.362	0.493
				(0.251)	(0.545)
$Timec_{ti}$ $XBMI_{ti}$(τ_{33})					0.004
					(0.005)
Covariance(τ_{01})		−1.328	−1.306	1.866	−5.687
		(0.371)	(0.369)	(1.662)	(23.649)
Covariance(τ_{02})				−5.583	−6.781
				(5.484)	(12.006)
Covariance(τ_{12})				−0.140+	0.585
				(0.078)	(1.087)
Covariance(τ_{13})					−0.074
					(0.113)
Covariance(τ_{23})					−0.030
					(0.047)
Model Fit					
N	4508	4508	4508	4462	4462
LL	−17515.280	−17218.927	−17216.041	−16929.270	−16918.370
AIC	35036.560	34449.854	34448.082	33884.540	33872.740
BIC	35055.801	34488.336	34499.391	33884.540	33988.000

括号内为标准误

+ $p<0.1$，* $p<0.05$，** $p<0.01$，*** $p<0.001$

注意，对于基于MLM的增长模型，结果整理的格式与标准的MLM模型有所不同，是按照所谓“初始状态”（截距）和“增长速率”（斜率）来结构化的，但各要素其实与MLM一致，只是呈现的形式不同而已。一般MLM结果也可以按照截距和斜率分开的格式进行整理。

7.3　曲线增长模型

线性增长模型，也就是假定个体发展是平滑的线性的。然而，实际中，个体的变化也可以是不连续或者非线性的。曲线增长模型属于非线性增长模型的一种，即结果变量随时间而变化的趋势是曲线的形式，可以用时间尺度变量的多项式函数（如二次方、三次方）模拟。这类曲线增长模型以预测变量的变化来实现，实际中还存在所谓“真正的非线性模型”（truly nonlinear model），即通过参数界定而实现的非线性模型，比如logistic或者negative exponential增长曲线模型，这类模型不满足动态一致性（dynamic consistency），即个体的与平均的轨迹函数形式不同（Singer＆Willett，2003：225-242）。

对于非连续（discontinuous）增长轨迹，如果可以假定不同阶段的变化可用直线模拟，即分段（piecewise）增长模型。对于更复杂的情况，可使用样条函数（spline function）来刻画增长曲线。一般来讲，在构建增长模型开始，需要依据理论假设，或从经验出发通过结果变量平均增长趋势图等形式，判断增长趋势是直线还是曲线的形式。更严谨的方式是结合假设，通过检验在模型中设定时间尺度变量

多项式函数的相应系数是否显著，判断个体变迁轨迹是线性还是非线性的趋势。多数情况下，尤其是观测点较少时，函数设定过于复杂，模型可能无法完全拟合，线性增长模型更为可取，至少可以作为复杂过程的较好的近似（Raudenbush＆Bryk，2002：163）。因为线性模型更为简约，有时曲线增长模型可以通过对结果变量或者水平-1时间尺度变量的线性变换，使得非线性增长模型变为线性增长模型。

曲线增长模型建模过程与线性增长模型基本一致。不过有时需要注意高次方多项式增长模型中的共线性问题。考虑带有3个时间点的曲线模型，t ＝0，1和2，t^2 ＝0，1和4，两个变量几乎完全相关。为了避免这一点，实践中有时进行中心化。让$(t-\bar{t})=-1,0$ 和1，$(t-\bar{t})^2=1,0$和1。如果3个时间点有同样数量的观测值，中心化处理就会去掉线性和二次方项之间的共线性（Hedeker＆Gibbons，2006：86）。

一般来讲，调查波次为单数次，可以中间值作为中心点对时间变量进行中心化。双数次没有观测的中间值，可用所有时间点除以波数得到理论上的中间值。这里的数据测量次数为单次，可取8作为中间值，虽然由于非等间隔测量，计算值是（0＋2＋4＋8＋11＋15＋17）／7＝8.14。中心化之前时间变量取值为0、2、4、8、11、15、17，平方项取值为0、4、16、64、121、225、289，相关系数为0.9661。经过中心化处理后，时间变量取值为－8、－6、－4、0、3、7、9，平方项取值为64、36、16、0、9、49、81，相关系数为0.1170。经过中心化之后，模型截距项的意义发生了改变。未中心化时，截距代表了第一个时间点时的差异。中心化之后，截距则代表了中心点时的差异。[2]

由于血压的变化不太符合曲线增长的趋势，下面以儿童身高为例，数据同样来自CHNS，为1989年0岁儿童队列升高的测量。

从图7-4来看，儿童身高增长符合二次曲线趋势。因此，这里构建带有性别非时变变量的曲线增长模型，相应的M*plus* 代码见框7-15。

框7-15　儿童身高的曲线增长模型

```
TITLE: Two level curve growth model;
DATA:
  FILE IS chnsheight.dat;
VARIABLE:
  NAMES＝id wave height age gender;
  USEVARIABLES＝id height gender age agesq;
  Missing are all（－9999）;
  WITHIN＝age agesq;
  BETWEEN＝gender;
  CLUSTER＝id;
DEFINE: agesq＝age*

age;
ANALYSIS: TYPE＝TWOLEVEL RANDOM;
          ESTIMATOR＝ML;
MODEL:
    %WITHIN%
    s1|height ON age;
```

```
s2 | height ON agesq;
%BETWEEN%
s1 s2 height ON gender;
s1 with s2;
s1 with height;
s2 with height;
```

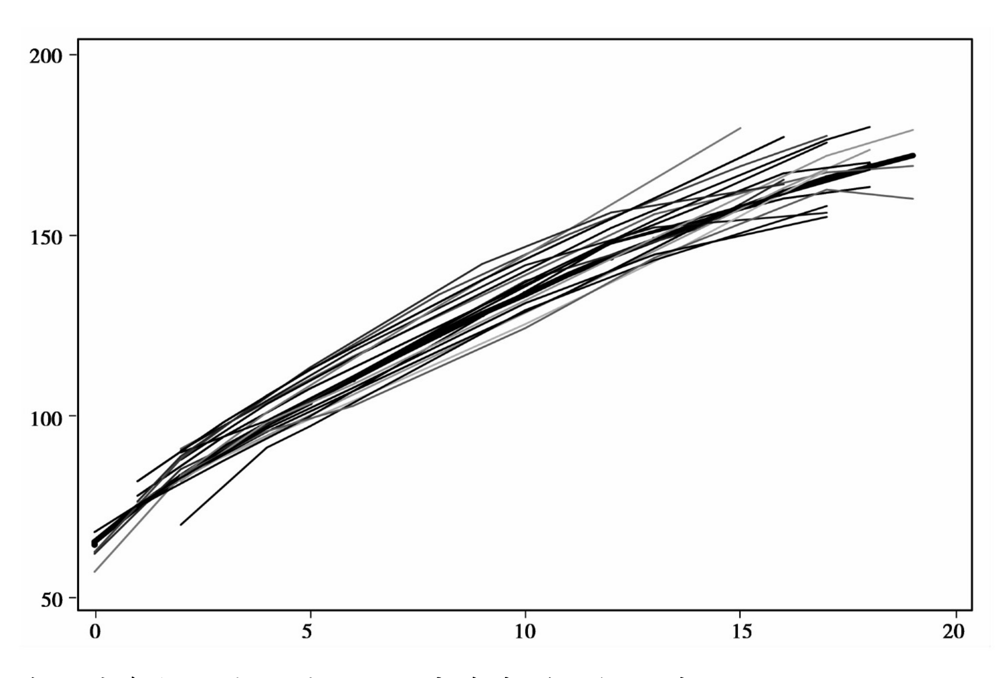


图7-4　采用非参数方法平滑化的儿童身高增长轨迹差异

（调查波次7次的前20位个体）

表7-2为相关模型输出结果整理而成的回归表。上述代码执行的结果见表7-2模型（4）。从结果来看，所考虑的几个模型中，二次曲线增长模型对儿童升高的拟合相对更好。

表7-2　曲线增长模型：CHNS1989年0岁组队列身高

	(1)	(2)	(3)	(4)
Fixed Effect				
Initial status, π_{0i}				
Intercepts(β_{00})	71.856***	71.050***	67.815***	66.681***
	(0.412)	(0.593)	(0.422)	(0.610)
gender(β_{01})		1.519+		2.039*
		(0.815)		(0.839)
Rate of change for age_{ti}, π_{1i}				
Intercepts(β_{10})	6.202***	6.124***	8.627	8.997***
	(0.065)	(0.100)	(0.137)	(0.205)
gender(β_{11})		0.110		−0.623*
		(0.123)		(0.270)
Rate of change for age_{ti}^2, π_{1i}				
Intercepts(β_{20})			−0.168***	−0.206***
			(0.009)	(0.013)
gender(β_{11})				0.062
				(0.016)
Variance Components				
Within-person(σ^2)	39.922***	39.983***	17.769***	18.318***
	(2.926)	(2.938)	(1.496)	(1.518)
Initial status(τ_{00})	5.708	5.335	14.798***	14.041***
	(3.872)	(3.804)	(3.602)	(3.518)
Rate of change(τ_{11})	0.173*	0.160*	0.907*	0.824*
	(0.075)	(0.075)	(0.366)	(0.344)
Covariance(τ_{01})	0.235	0.444	−0.683	−0.572
	(0.468)	(0.426)	(0.946)	(0.895)
Model Fit				
N	733	733	733	733
LL	−2497.901	−2494.058	−2328.449	−2316.092
AIC	5007.802	5004.115	4676.898	4658.184
BIC	5035.384	5040.892	4722.870	4717.946

括号内为标准误

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

处理增长模型（以及多水平回归模型），一般还有一个重要主题是进行残差协方差结构的设定问题。选择不同的协方差结构不会对固定效应参数估计产生显著的影响，但是选择合适的协方差结构有助于提高参数估计的精度。设定残差协方差矩阵结构在一些使用复合模型

框架处理历时数据的软件中较为便捷。[3] *Mplus* 中对于残差协方差结构并没有现成的命令参数，需要特制性设定，基本思路是将残差表达为因子。为此，*Mplus* 需要使用宽型数据或部分宽型数据。由于这种方法已经属于在结构方程模型框架下建模，需以潜增长模型为基础。[4]

7.4　本章小结

在进行历时数据分析时，与潜增长模型（见第4 、第5章 ）的思路不同，也可以将不同轮次的测量视作嵌套在个体内部，从而用多水平回归模型进行建模。

本章介绍了多水平回归模型框架下基础的线性增长模型和曲线增长模型，包括无条件均值模型、无条件增长模型，以及纳入非时变和时变协变量的多水平增长模型。

多水平增长模型可作为潜增长模型的一个替代性选择。不过在以潜变量建模见长的*Mplus* 中，对于一些较复杂的历时数据分析，推荐优先使用潜增长模型。

※推荐阅读※

Singer＆Willett（2003）是多水平框架下处理历时数据的经典之作。上一章所提及的有关多水平回归模型的著作，多数也含专门章节介绍多水平框架下增长模型的处理，可作为参考。其他文献参见

Singer＆Martin（1998）、Guo＆John（2004）、Steele（2008）等。中文介绍文献如宋时歌、陈华珊（2005）。

基于MLM的增长模型经典研究如Huttenlocher et al（1991 ）、Raudenbush & Chan （ 1993 ） 、 Johnson （ 2002 ） 、 Sampson & Sharkey（2008）等。

---

(1) 多次观测数据可以得到初始状态与真实变化之间相关一致性的估计。在前-后测的研究中，由于前测时测量误差和观测的变化值有负相关，会导致初始状态与增长率之间的虚假负相关（Raudenbush＆Bryk，2002：167）。

(2) 使用正交多项式（orthogonal polynomials）的时间量度也是避免多项式共线性问题的一种方法。而且，使用原始的时间测度，高次方多项式的回归系数估计会越来越困难，系数及其标准误会越来越小。使用正交多项式则可以消除这类问题（Hedeker＆Gibbons，2006：86）。

(3) 如SAS中，在TYPE选项可以选择特定类型残差协方差结构，如UN、CS、AR（1）、TOEP等类型。Stata中包括independent、exchangeable、ar＼＃、ma＼＃、unstructured、banded、toeplitz＼＃、exponential等类型。

(4) 也可以借助R中的M*plus* Automation包中的M*plus* Rcov命令，间接辅助对M*plus* 残差协方差结构进行设定。

8　多水平结构方程模型

8.4.2　多水平结构方程模型

8.5　本章小结

8.1　多水平结构方程模型的基本框架

8.1.1　多水平结构方程模型的定义

嵌套或聚类数据结构破坏了许多统计模型所需的独立性假定，多水平回归模型可以用来解决独立性假设破坏的问题。然而多水平回归模型存在局限，如没有潜变量（测量模型），不能处理测量误差；没有全局性拟合优度指标；模型较为简单，未能考虑组内和组间结构的复杂性，等等。

传统的路径分析虽然可以厘清变量之间复杂的因果关系，但没有注意到存在的嵌套结构。比如，学生阅读得分呈现较强的组内相关，利用空模型计算ICC＝0.203，提示需要关注客观存在的嵌套结构。如果我们想分析学生阅读得分在学校之间的差异多大程度上可由学校层面的变量来解释，或者进一步检验学校层面因素对于学生阅读得分校际差异的直接和间接效应，就需要纳入对数据嵌套结构的考虑，引入多水平路径分析。当然，路径分析没有考虑测量误差。如果同时满足上述这些条件，就需要引入标准的多水平结构方程模型（multilevel structural equation model，以下简称MSEM）的框架。

实际上，多元回归、路径分析、验证性因子分析（CFA）、结构方程（SEM）、多水平回归模型（以下简称MLM）、潜发展模型等多种常见模型，都可以统一到MSEM的框架之下，视为MSEM的特例（Muthén＆Asparouhov，2011）。

作为一个通用分析框架，MSEM结合了MLM和SEM的优势。MSEM适用于嵌套数据模型，且通过各构念多观测指标测量并使用潜聚合方式形成高水平构念的设置，校正了MLM在进行情景分析时存在的测量误差和抽样误差，在方法上更具优势。

构架MSEM有两种途径：整合潜变量和测量误差到多水平回归模型，如HLM、MLwiN；整合多水平结构包括随机截距和斜率到SEM，如M*plus*、GLLAMM。由于本书以介绍M*plus*应用为主旨，所以这里采用M*plus*对结构方程模型定义的体例。根据Muthén & Asparouhov（2008）的定义，单水平结构方程模型测量方程为[(1)]：

$$Y_i = \nu + \lambda \eta_i + KX_i + \varepsilon_i \quad （8.1）$$

其中，i 为个体指示；Y_i 为$p \times 1$维观测变量向量；ν 为$p \times 1$的截距向量；η_i 为$m \times 1$维潜变量向量；λ 为$p \times m$ 维对m 个潜变量的因子载荷矩阵；η_i 为m 个潜变量的$m \times 1$维向量；K 为q 个外生协变量X_i 对p 个观测变量Y_i 回归斜率构成的$p \times q$ 维矩阵；X_i 是q 个观测协变量构成的$q \times 1$维向量；ε_i 为$p \times 1$维残差向量，假定服从均值为0、协方差矩阵为Θ的多元正态分布。测量模型即验证性因子分析，表达了观测变量（Y_i）作为潜变量（η_i）、观测的协变量（X_i）和残差（ε_i）的函数。

结构方程设定了潜变量之间的关系模式：

$$\eta_i = \alpha + B \eta_i + \Gamma X_i + \zeta_i \quad （8.2）$$

其中，α 为$m\times 1$维截距项向量；B 为设定潜变量之间关系的$m\times m$维结构回归系数矩阵；Γ 为潜变量对外生变量回归斜率系数的$m\times q$维矩阵；η_i 同前；ζ_i 为$m\times 1$维潜变量回归残差向量。ζ_i 中的残差假定为均值为0、协方差矩阵为Ψ的多元正态分布。结构方程可以看作潜变量（η_i）之间及与外生观测协变量（X_i）和残差（ζ_i）函数关系的表达。

MSEM可以视作SEM的拓展，为了体现组群水平随机效应，允许某些系数矩阵在不同组群之间变化（Muthén & Asparouhov，2008；Preacher et al，2011；2010）。

测量模型：

$$Y_{ij}=\nu_j+\lambda_j\eta_{ij}+K_jX_{ij}+\varepsilon_{ij} \quad (8.3)$$

组内结构模型（within-group structural model）：

$$\eta_{ij}=\alpha_j+B_j\eta_{ij}+\Gamma_jX_{ij}+\zeta_{ij} \quad (8.4)$$

组间结构模型（between-group structural model）：

$$\eta_j=\mu+\beta\eta_j+\gamma X_j+\zeta_j \quad (8.5)$$

MSEM中，Θ和Ψ中元素假定不随组群而变化。含有模型参数的矩阵，包括ν_j、α_j、λ_j、K_j、B_j、Γ_j既可固定，也可在组间变化。其中，j 为组群指示标记，代表该参数随组群而变化。

需要注意的是，式8.5中，η_j与8.3中η_{ij}不同。η_j包括所有r个随机效应，堆栈式8.3和式8.4中所有下标为j参数矩阵的随机元素（即ν_j、λ_j、K_j、α_j、B_j、Γ_j），这是MSEM超越SEM的最主要的创新（Preacher，2015）。其形如：

$$\boldsymbol{\eta}_j = \begin{bmatrix} \text{vec}\{\boldsymbol{\nu}_j\} \\ \text{vec}\{\boldsymbol{\alpha}_j\} \\ \text{vec}\{\boldsymbol{\Lambda}_j\} \\ \text{vec}\{\boldsymbol{K}_j\} \\ \text{vec}\{\boldsymbol{B}_j\} \\ \text{vec}\{\boldsymbol{\Gamma}_j\} \end{bmatrix} \quad (8.6)$$

X_j为所有组群水平协变量的向量，假定为s维。X_j与X_{ij}也不一样，X_j是所有组群层面协变量堆栈而成的s维向量。

向量$\boldsymbol{\mu}$（$r\times1$）、矩阵$\boldsymbol{\beta}$（$r\times r$）和矩阵$\boldsymbol{\gamma}$（$r\times s$）包含所有估计的固定效应。其中，$\boldsymbol{\mu}$为$r\times1$维固定效应向量，包含随机效应分布的均值、组间结构方程的截距；$\boldsymbol{\beta}$为$r\times r$维向量，包含潜变量及随机截距和斜率彼此之间的关系的结构回归斜率系数；$\boldsymbol{\gamma}$为$r\times s$矩阵，包含η_j中随机效应对于s各组群水平外生预测变量的回归斜率。ζ_j中组群水平残差服从均值为0、协方差矩阵为$\boldsymbol{\Psi}$的多元正态分布（Muthén & Asparouhov，2008；Preacher et al，2010）。

上述Muthén＆Asparouhov（2008）的定义方式，赋予MSEM以更大的灵活性。包含变量取值的矩阵（包括观测的Y_{ij}、X_{ij}和不能观测的η_{ij}、ε_{ij}、ζ_{ij}）可在两个水平变化。后一种矩阵中的元素可以严格限定为组内变量，也可以严格限定为组间变量，或者是在两个水平均有变化的变量（Boulton，2011）。同时，与另一通用潜变量建模程式GLLAMM相比，M*plus* 具有计算效率更高、允许随机斜率等优势。

刘东等（2012）总结了相比于MLM使用MSEM的优势。包括：（1）结构方程模型或路径模型可以同时检验多重关系，并展示整个模型的拟合程度。这样通过模型对比，来提供检验假设模型与其他备选模型的机会；（2）在同一模型中检验多重关系，可以减少对模型参数和标准误估计的偏差。当进行复杂多层次关系估计时，这种情况更是如此；（3）当一个或多个变量中存在缺失值时，同时估计多重关系也有利于对缺失值进行估计，促进模型的构建；（4）在模型估计时，结构方程模型（即潜变量模型）考虑了测量误差，这样可以更为准确地估计变量之间的关系。

8.1.2　多水平潜协变量方法与双潜多水平模型

进行多水平模型的估计，基于MLM与基于MSEM的思路存在一个重要差异，即如何看待和处理“聚合”而成的水平-2的构念。

在传统的MLM分析中，组群水平的构念经常来自于将个体层面数据聚合到组群层面构建。然而这样的处理方式经常缺乏信度，特别是各

组观测数较少、ICC较小时，结果导致情境效应的估计偏差（Raudenbush et al，1991；Snijders＆Bosker，1999；Bliese，2000）。这一点从相应信度估计公式即可以看出：

$$\text{L2 Realiability}(\bar{X}_{.j}) = \frac{n \times \text{ICC}}{1 + (n - 1) \times \text{ICC}} \tag{8.7}$$

其中，n是一般的组群规模；ICC为组内相关系数。

按照这种方式生成“情境变量”，实际上假定了基于水平-1变量聚合而得的水平-2变量没有测量误差，然而协变量往往带有测量误差，真实值是不可观测或潜在的（Rabe-Hesketh et al，2003）。

对此，Muthén、Asparouhov和Lüdtke等（Asparouhov＆Muthén，2006；Lüdtke et al，2008；Muthén & Asparouhov，2011）以Raudenbush＆Bryk（2002：140）提到的一个多水平回归模型为例，指出MLM在处理协变量时存在偏差：

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \bar{X}_{.j}) + \varepsilon_{ij} \tag{8.8}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\bar{X}_{.j} + u_{0j} \tag{8.9}$$

$$\beta_{1j} = \gamma_{10} \tag{8.10}$$

将式8.9、式8.10代入式8.8并整理可得：

$$\begin{aligned} Y_{ij} &= \gamma_{00} + \gamma_{01}\bar{X}_{.j} + \gamma_{10}(X_{ij} - \bar{X}_{.j}) + u_{0j} + \varepsilon_{ij} \\ &= \gamma_{00} + (\gamma_{01} - \gamma_{10})\bar{X}_{.j} + \gamma_{10}X_{ij} + u_{0j} + \varepsilon_{ij} \end{aligned} \tag{8.11}$$

定义“情境效应”（contextual effect）为：

$$\beta_c = \gamma_{01} - \gamma_{10} \quad （8.12）$$

$X_{.j}$ 可视作对水平-2未被直接观测构念的一个估计。实际上，协变量 $X_{ij} - \bar{X}_{.j}$ 和 $\bar{X}_{.j}$ 可以被看作潜协变量的代理（Asparouhov & Muthén，2006；Muthén＆Asparouhov，2011）：[2]

$$X_{ij} - \bar{X}_{.j} \approx X_{ijw} \quad (8.13)$$

$$\bar{X}_{.j} \approx X_{jb} \quad (8.14)$$

在式8.13、式8.14中，下标 w 和 b 分别指代组内和组间部分。基于此，在M*plus* 中处理多水平结构方程模型的思路即所谓“多水平潜协变量”（multilevel latent covariate，MLC）模型的思路。与之相对，传统的MLM方法，将组均值作为显变量，而不是将之作为不可观测的潜在构念，被称为“多水平显协变量”（multilevel manifest covariate，MMC）模型的方式。

潜协变量方法的基础就是各变量均可区分出不可观测部分，考虑作为潜变量。比如对于变量 X_{ij} ，可分解为组均值成分和个体对组均值成分的偏差，也就是组间和组内两个独立的部分（Muthén＆Asparouhov，2011）：[3]

$$X_{ij} = X_{jb} + X_{ijw} \quad （8.15）$$ [4]

同样，Y_{ij} 也可以分解为组间和组内两部分：

$$Y_{ij} = Y_{jb} + Y_{ijw} \quad （8.16）$$

其中，$E(X_{ijw})=0$，$E(Y_{ijw})=0$；$Cov(X_{jb}, X_{ijw})=0$，$Cov(Y_{jb}, Y_{ijw})=0$。

X_{ij}和Y_{ij}是观测值，Y_{jb}、X_{jb}、Y_{ijw}、X_{ijw}均属于不可观测的随机成分。由此，使用潜协变量方法，多水平回归模型可以表述为：

$$Y_{ijw} = \beta_w X_{ijw} + \varepsilon_{ij} \quad (8.17)$$

$$Y_{jb} = \mu + \beta_b X_{jb} + \varepsilon_j \quad (8.18)$$

将式8.16、式8.17代入式8.18，整理可得：

$$\begin{aligned} Y_{ij} &= Y_{jb} + \beta_w X_{ijw} + \varepsilon_{ij} \\ &= \mu + \beta_b X_{jb} + \beta_w (X_{ij} - X_{jb}) + \varepsilon_j + \varepsilon_{ij} \\ &= \mu + (\beta_b - \beta_w) X_{jb} + \beta_w X_{ij} + \varepsilon_j + \varepsilon_{ij} \end{aligned} \quad (8.19)$$

据此，“情境效应”（contextual effect）即为：

$$\beta_c = \beta_b - \beta_w \quad (8.20)$$

Asparouhov＆Muthén（2006）和Lüdtke et al.（2008）证明，在进行多水平回归模型分析时，与潜协变量方法相比，观测协变量方法在估计水平-1 γ_{10} 时不存在偏误，但在估计水平-2斜率 γ_{01} 时引入了偏误：

$$E(\hat{\gamma}_{01}) - \beta_b = (\beta_w - \beta_b) \frac{1}{n} \frac{1 - ICC}{ICC + (1 - ICC)/n} \quad (8.21)$$

随着组群规模或ICC减小，估计偏误会增加。类似地，也可以证明，在估计情境效应时，观测协变量方法$\hat{\gamma}_{01}-\hat{\gamma}_{10}$对潜协变量方法$\beta_b - \beta_W$的估计也存在偏误。

$$E(\hat{\gamma}_{01}-\hat{\gamma}_{01})-(\beta_b-\beta_w) = (\beta_w-\beta_b)\frac{1}{n}\frac{1-ICC}{ICC+(1-ICC)/n} \quad (8.22)$$

该结果暗示，如果$\beta_W < \beta_b$，则情境效应被低估，相反则被高估。模拟研究的结果也支持数学推导的结论（Lüdtke et al，2008）。

这种方法的使用不限于估计情境效应，实际上，按照Lüdtke et al（2008；2011）和Preacher et al（2011；2016）、Preacher（2015）的概括，MSEM建模是一种“分解优先”（decomposed-first）思路。在MSEM框架下，水平-1（组内/组间两用）观测变量自动分解为组内（W）和组间（B）不同部分，从而区分出组内和组间效应。[(5)]这种策略的优势在于，相对于MLM方法，将水平-1变量组间部分作为潜变量（潜组均值，实际即随机截距），从而减少了由于在水平-2采用手工计算组均值带来的偏误。同时，对组群水平系数估计进行校正。M*plus*所运用估计方法（默认基于EM算法的MLR）满足一致性和渐进有效性（Asparouhov & Muthén，2006；Lüdtke et al，2008）。[(6)]

不过，MSEM带有更多需要估计的参数，以牺牲效率为代价，可能会减少侦测间接效应的功效，特别是在ICC较小的情况下（Preacher et al，2011；Preacher，2015）。Lüdtke et al（2008）指出，虽然

相对而言MLC方式估计的情境效应更有效和一致，因此优于MMC方法，然而MLC方法对水平-2样本量的要求更高。如果水平-1和水平-2样本量处于中等水平，ICC需要非常大才推荐使用MLC。如果各组样本量都足够大时，二者的结构相近。如果水平-1样本量和入样比例都比较小，两种方法结果差异就会非常不同。平均而言，真值应落在两种方法估计的结果的区间之内。因此，推荐两种方法同时运用，如果结果差异较大，需要小心推论。

另外，如果水平-1的指标本身就是测量水平-2的构念，那么使用MLC的方式更适合。Lüdtke et al（2008）具体区分了聚合构念的不同类型。依照形成性（formative）和反映性（reflective）构念的区分，聚合也可以分为形成性和反映性两种类型。在反映性聚合类型中，假定了个体水平数据和组群水平构念的同构关系。组群水平构念假定是由个体水平对应的构念测量的，参照点是组群水平构念，水平-1指标反映性聚合至水平-2构念。举例如通过对学生的调查测量“校风”。这种情况下，应用MLC方式更适合，即使抽样比例比较大。在形成性聚合类型中，聚合过程假定组群水平变量只是一个界定充分、被聚合到水平-2的水平-1构念的指标。聚合基于离散个体的不同属性，构念主要的目的是反映水平-1个体的差异，参照点是水平-1构念，比如学生的家庭社会经济地位。此种情况下，如果抽样比例较高，样本量和ICC较大，MMC方法可以使用。当抽样比例较小，而水平-1和水平-2样本量较大时，考虑使用MLC方法。其他情况下，应该两种方法都使用，若结果差异较大，研究者应谨慎推论。另外，如果水平-2变量对应的总体均值或比例已知（比如，学校性别比），并不存在信度问题，应采用MMC方法。

虽然上述MMC和MLC不同思路的公式看起来较为复杂，但在M*plus*中二者的代码区分起来却相当直观。从代码书写规范的角度来讲，二者的区别就在于，对于特定变量，是否在变量声明部分（VARIABLE ）设定水平-1变量所属的层次，即指定变量属于WITHIN还是BETWEEN层次，以及在模型部分（MODEL）是否分别在组内（WITHIN）和组间（BETWEEN）设定相应的变量名。如果在VARIABLE部分不具体指定变量属于哪一层次，也就是属于所谓“组内／组间两用”变量（Asparouhov＆Muthén，2006），那么按照MLC的思路，M*plus* 默认将该变量分解为组内和组间两个不相关潜协变量部分，并同时在MODEL部分组内和组间两个层面参与建模。

传统的MLM模型使用的是MMC方式，只有水平-1结果变量y的变量名同时出现在WITHIN和BETWEEN两个部分，属于作为“组间／组内两用”，其他协变量只属于特定的某个层次，即所谓“组内专用”（within-only）或“组间专用”（between-only）的类型。由于水平-1的协变量在WITHIN中定义，因此不会在水平-2方程中出现。所以，在MLM模型中，如果水平-2需要水平-1对应的变量，需要手动将水平-1相应变量聚合（也就是取组均值或比例）到水平-2，然后指定其属于BETWEEN层面，才能参与水平-2模型建构。

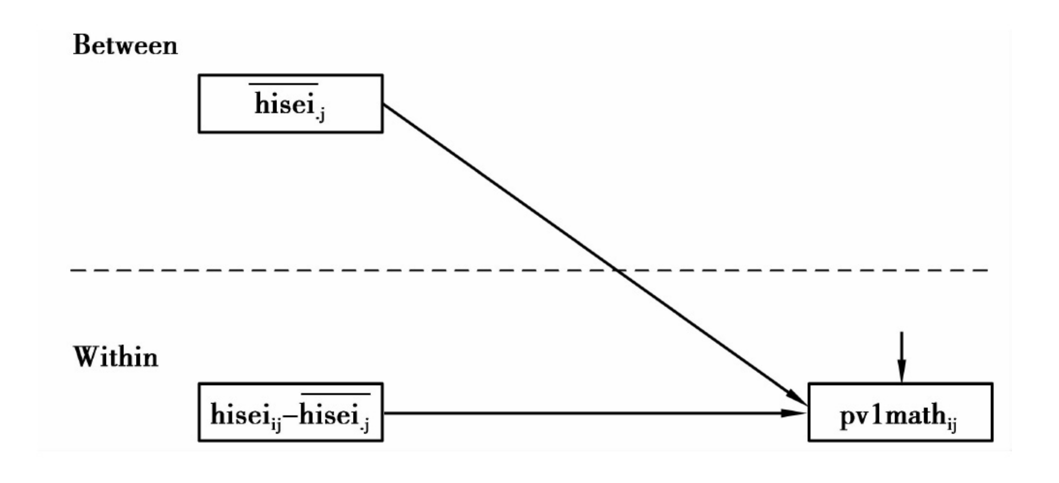


图8-1　MMC方法

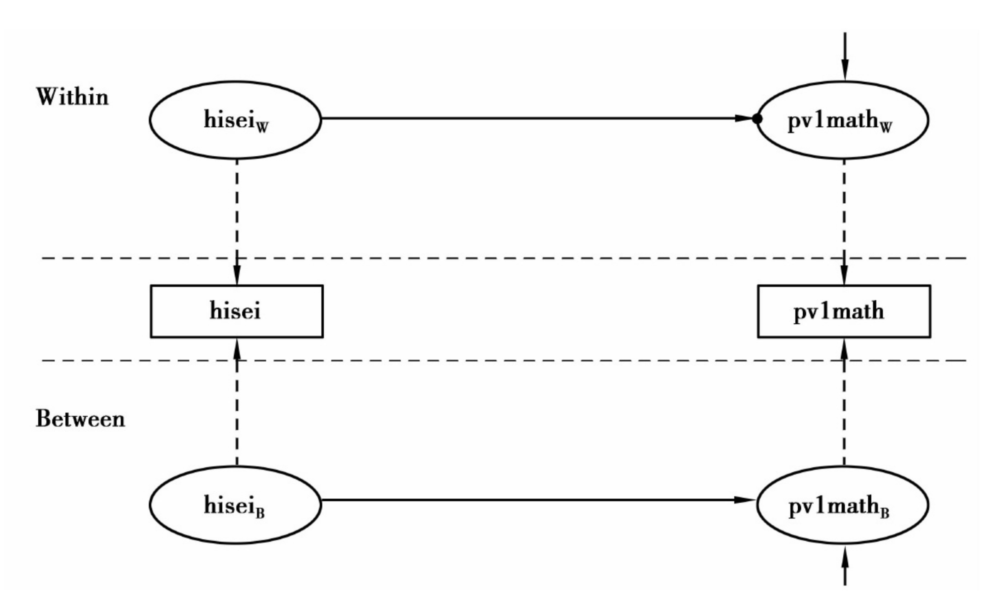


图8-2　MLC方法

这里不妨针对PISA数据中的相应变量分别以MMC和MLC建模，通过对代码和结果进行比较，增进对二者建模思路差异的认识。

框8-1　利用MMC方法进行情境效应的分析

```
Title: Multilevel Manifest Covariate Approach;
Data:
  File is pisa_MMC_MLC.dat;
Variable:
  Names are schoolid hisei pv1math grouphisei;
  USEVARIABLES＝schoolid hisei pv1math grouphisei;
  Missing are all（－9999）;
  Cluster＝schoolid;
  Within＝hisei;
```

```
  Between＝grouphisei;
DEFINE: CENTER hisei（GROUPMEAN）;
ANALYSIS:
  TYPE＝TWOLEVEL;
MODEL:
  %WITHIN%
    pv1math on hisei（gamma10）;
  %BETWEEN%
    pv1math on grouphisei（gamma01）;
MODEL CONSTRAINT:
  NEW（betac）;
  betac＝gamma01－gamma10;
```

使用MMC方法建模时，纳入水平-1方程的家庭社会经济地位进行了组中心化处理（这里未预处理，而是在*Mplus*代码DEFINE部分进行处理），水平-2方程纳入的则是预处理的不同学校学生平均的家庭社会经济地位变量（grouphisei），同属显变量。在变量声明部分，分别指定了二者所属层次。而在运用MLC方法的模型中，参与建模的是家庭社会经济地位的原始变量。变量声明部分，没有指定变量隶属的层次。模型部分，在WITHIN层次，结果变量和预测变量分别指向Y_{ijw}和X_{ijw}，而在BETWEEN层面，则分别指向Y_{jb}和X_{jb}。

框8-2　利用MLC方法进行情境效应的分析

```
Title: Multilevel Latent Covariate Approach;
Data:
  File is pisa_MMC_MLC.dat;
Variable:
  Names are schoolid hisei pv1math grouphisei;
  USEVARIABLES＝schoolid hisei pv1math;
  Missing are all（－9999）;
  Cluster＝schoolid;
  Within＝;
  Between＝;
ANALYSIS:
  TYPE＝TWOLEVEL;
MODEL:
  %WITHIN%
    pv1math on hisei（gamma10）;
  %BETWEEN%
    pv1math on hisei（gamma01）;
MODEL CONSTRAINT:
  NEW（betac）;
  betac＝gamma01-gamma10;
```

默认最大似然估计（MLR），两种方式的输出结果见表8-1。

表8-1　基于MMC和MLC估计情境效应分析

	MMC	MLC
Fixed Effect		
Intercepts(γ_{00})	226.972 *** (19.583)	190.698 *** (23.278)
hisei-group_hisei(γ_{10})	1.102 *** (0.080)	
group_hisei(γ_{01})	5.065 *** (0.377)	
hisei_w(γ_{10})		1.102 *** (0.080)
hisei_b(γ_{01})		5.763 *** (0.449)
Random Effect		
Intercepts(τ_{00})	773.813 *** (102.462)	644.351 *** (104.310)

续表

	MMC	MLC
Residual(σ^2)	5428.988 *** (122.939)	5428.988 *** (122.939)
Model Fit		
N	4763	4763
LL	-27365.911	-47075.888
AIC	54741.823	94161.777
BIC	54774.166	94194.120

括号内为稳健标准误

+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001

从结果来看，MMC和MLC方法在WITHIN部分的系数是一致的，γ_{10} 均是1.102。然而，BETWEEN部分有较大差异，MMC方法 γ_{01} ＝5.065，而MLC方法 γ_{01} ＝5.763。计算情境效应（β_c），MMC：5.065－1.102＝3.963，MLC：5.763－1.102＝4.661（代码中也给出了

通过增加MODEL CONSTRAINT自动输出情境效应的命令）。如果MLC的估计更接近真实值的话，那么从结果来看MMC低估了情境效应。

不过MLC针对的是单指标观测变量，虽然可以控制抽样误差但并没有考虑测量误差（Lüdtke et al，2011）。通过强调各构念多观测指标测量并使用潜聚合方式形成水平-2构念，Marsh、Lüdtke等（Marsh et al，2009；Lüdtke et al，2011；Marsh et al，2012；Morin et al，2014）将应用于情境效应分析中的MSEM称作“双潜多水平模型”（doubly latent multilevel model）。因为这样的设置实现了同时对水平-1与水平-2由于“题项抽样”（sampling items）造成的测量误差和由于在水平-1特征聚合构建水平-2构念过程中“个体抽样”（sampling of persons）存在的抽样误差的校正。按照这样的标准，传统MLM则属于双显模型（doubly manifest model），特点是显测量、显聚合，两种误差均未考虑。如果构念采取的是单指标而非多指标测量，但通过潜聚合控制了抽样误差，属于显-潜模型（manifest-latent model）。MLC即属于这种情况（Lüdtke et al，2011）。一些组群层面指标构建研究中常用的将各指标进行聚合后在组间进行单水平因子分析的处理方式属于典型的潜-显模型，亦即通过多指标控制了测量误差但没有通过潜聚合控制抽样误差。潜-显模型与显-潜模型同属部分校正（partial correction）模型。标准的MSEM属于双潜模型，各观测指标基于MLC方法进行组间／组内分解（除非严格属于特定水平的专用变量），同时各水平潜在构念基于（分解后）多指标进行估计。

上述方式赋予MSEM相对于MLM所不具备的优势。一方面，MSEM实现了对测量误差和抽样误差双重校正，减少了由于在水平-2采用聚合组均值带来的偏误。同时，组群层面系数估计也得到校正。另一方面，

MSEM在进行多水平分析时层次分明、更具灵活性。各层可以只考虑本层要素的方差，不同水平之间互相独立。比如，多水平验证性因子分析（以下简称MCFA）中，可允许不同水平设定具有差异的因子结构。多水平中介或调节效应分析中可方便纳入不同层面协变量进行“跨水平”中介效应分析，而不会产生传统MLM进行多水平中介效应分析（除非进行必要处理否则）会出现的组内／组间效应混淆的问题（Krull＆MacKinnon，2001；Zhang et al，2009；Preacher et al，2010；Preacher et al，2011；Li＆Beretvas，2013；Preacher et al，2016）（参见本书第9章），相对于MLM估计更为简单也更加准确。

8.1.3　估计方法

（1）Muthén“伪平衡法”

在Muthén（Muthén，1989；1990；Muthén＆Satorra，1995）的早期讨论多水平协方差结构模型（multilevel covariance structure model）的论文中，为了处理样本非均衡（即各组样本量不同）的问题，采取一个变通的方式，设置了c参数（接近于平均的样本规模）（详见本章第2节）。这种解决方式不是充分信息最大似然估计（FIML），而是所谓有限信息最大似然估计（LIML），或称为“伪平衡法”（pseudo-balanced soulution）（McDonald，1994），或MUML（Muthén's ML-based maximum likelihood estimation）（Muthén，1989；1990；1994）。如果两层样本量都足够大，各组群样本量相差不是很悬殊，伪平衡法的估计也将与充分信息最大似然估计（FIML）接近（Muthén，1989；1990）。

近来有模拟研究（Yuan＆Hayashi，2005）则指出，大样本并无助于改善不平衡数据的精确性。对于严重不平衡数据，MUML会产生有偏的标准误和显著性检验，且不会随着样本量增加而消除（Hox，2010）。

目前在M*plus* 中MUML法仍可以使用，但实际上有更好的方法，如MLR、WLSM（V）等（Asparouhov & Muthen，2003；Hox et al，2010）。新方法一般采取EM算法，并不特别要求分别估计组间和组内协方差矩阵（所谓联立估计，simultaneous approach），有的允许有缺失值和随机斜率，响应变量也可以是类型变量。

（2）充分信息最大似然估计

为了处理非均衡数据和缺失值的问题，FIML被引入SEM估计。非平衡数据，视为一种不完整数据（incomplete data），适用FIML估计方法（Arbuckle，1996；Neale，2000；Bauer，2003；Bentler & Liang，2003；Liang＆Bentler，2004）。[7]

目前M*plus* 、GLLAMM、Lisrel等程式都可以使用FIML处理多水平结构方程模型。而且，对于组间模型（between-groups model），两阶段法（包括MUML和WLS）只能包括随机截距，而FIML还可以将随机斜率整合进来（Mehta＆Neale，2005），也就是允许因子负荷跨层随机变动。[8]

Asparouhov＆Muthén（2003）提出了3种以EM为基础的最大似然法估计：ML、MLF和MLR。3种估计方法的区别主要在于标准误的估计。模

拟研究表明，MLR作为稳健的充分信息ML估计，不需要正态性假定，给出非正态稳健标准误估计，因此尤其对于非正态变量分布，具有更好的表现。MLR目前是M*plus* 针对多水平潜变量模型默认的估计方法。

（3）加权最小二乘法

对于类别变量和非正态数据，ML法估计需要高维数字积分，计算量很大。基于Muthén（1984）提出的WLS方法（weighted least squares estimation），Asparouhov & Muthén（2007）提出WLSM（weighted least squares-mean-adjusted estimator）估计方法，不仅可以完成对类别变量、删节、正态分布连续变量及其组合的多水平潜变量模型的估计，而且可带有随机截距，协变量数量任意。

WLS法使用样本方差协方差矩阵S_B 和S_W 作为权重矩阵，得到校正的卡方和标准误估计。标准的WLS进行估计时，构造所有待估参数的渐进协方差作为权重矩阵。随着因子和样本数增加，数值积分增加，运算量增大。尤其对于非约束模型，由于参数较多，矩阵会变得很大，甚至超过样本中组群个数，导致奇异矩阵，估计不能继续。因此除非组群规模特别大，一般采用矩阵的对角元素，即Diagonal WLS。

M*plus* 中有两种加权最小二乘方法：WLSM和WLSMV。区别在于卡方的计算。如果卡方采用均值校正（mean-adjusted），则为WLSM；如果采用均值-方差校正（mean-and variance adjusted），则为WLSMV。

需要注意的是，如果使用WLSMV（以及MLMV）进行估计，模型差异检验要通过SAVEDATA中设置DIFFTEST选项实现。

（4）贝叶斯法

对于某些样本量较小、计算量巨大、迭代收敛困难的MSEM模型，可以尝试贝叶斯估计（Baldwin＆Fellingham，2013；Depaoli，2015）。

Depaoli＆Clifton（2015）的模拟研究表明，由于MSEM的复杂性，有时最大似然法可能造成组间方差成分为负或参数估计有偏的问题。使用Markov chain Monte Carlo（MCMC）算法的贝叶斯法可以克服收敛问题，改善参数有偏估计。贝叶斯估计相关内容可参见本书第10章。

（5）不同方法比较

Hox（2010）指出，（基于频次框架的）几种方法相比较，伪平衡MUML法准确性较差，尤其是对组间模型的估计。模拟分析（Hox et al，2010）表明，WLS和FIML估计接近，准确性均较MUML要好。不过，尽管在数据违反多维正态分布时MLR估计要比ML更准确，但需要更大的样本规模。样本量比较小时，不推荐使用稳健方法。

Asparouhov＆Muthén（2007；2008）的模拟分析则表明，相比于EM和MCMC，WLSM（V）计算更有效率。如果潜变量是类型变量，WLSM（V）估计优于MLR。如果是连续的正态分布，则与MLR结果一致。几种常用方法及适用情况见表8-2。[9]

表8-2　常见估计方法及适用情况

估计方法	适用情况
MLR	M*plus* 默认设置，通用估计方法，适用于非正态数据
ML	适用于多元正态分布数据，没有未观测异质性
WLSM（V）	通用方法，模型复杂时运算速度更快，适用不完整数据或类别变量
WLS	只有当组群样本非常大时使用
BAYES	频次框架估计方法的替代性选择，适用于样本量小、迭代困难的情况

8.1.4　拟合指标

结构方程模型可以提供拟合指数，以评价模型与数据的拟合程度。相比于一般的多水平回归模型，SEM也具有这样的优势。[10]

评估模型拟合程度也是MSEM建模的重要内容。不过，由于TLI、CFI、RMSEA等常用指标源于单水平SEM，用以MSEM的评价存在一定问题。因为这些指标是针对模型整体而言的，未能考虑MSEM的多水平结构。所以实际中常常由于高水平样本量较小，这些笼统的指标难以有效评价高水平模型的设定问题。

表8-3　M*plus* 默认输出一般结构方程拟合指标

指　标	全　称	一般标准
CFI	比较拟合指数	取值为0～1，越接近0表示拟合越

	（comparative fit index）	差，越接近1表示拟合越好。CFI≥0.9模型拟合较好。
TLI	Tucker-Lewis指数（Tucker-Lewis index）	取值为0～1，越接近0表示拟合越差，越接近1表示拟合越好。TLI≥0.9模型拟合较好。
RMSEA	近似误差均方根（root-mean-square error of approximation）	RMSEA＝0模型完全拟合；RMSEA＜0.05模型接近拟合；0.05≤RMSEA≤0.08模型拟合合理；0.08＜RMSEA＜0.10模型拟合一般；RMSEA≥0.10模型拟合较差。
SRMR	均方根残差（standard root of the mean square residual）	SRMR＜0.1模型拟合较好。

故而，由此计算而得的指标，如CFI（comparative fit index）和RMSEA（root-mean-square error of approximation）等指数，可能主要反映的是水平-1的情况，对水平-2拟合差的问题缺乏敏感度（Yuan＆Bentler，2007；Ryu，2008；Ryu，2014）。即使这些指数显示拟合较差，也不清楚到底是在哪一层面较差，还是两个层面都表现欠佳。对此，出现了两种替代性方法：利用偏饱和模型分别评估不同水平模型（Ryu＆West，2009）以及通过分割多水平协方差结构为多个单水平协方差结构（Yuan＆Bentler，2007）。Ryu（2014）的模拟研究支持了上述两种方法相对于传统方法的优越性。

水平特定（level-specific ）的拟合评价对于鉴别MSEM设定问题更具优势（Schermellehengel et al，2014）。对于MSEM，目前M*plus*

默认输出结果中，SRMR是分水平输出的，是唯一可以分别用以评价不同水平模型的拟合程度的指标。其他拟合指标非水平特定。[(11)]

进一步的模拟分析（Hsu et al，2015）表明，对于MSEM，CFI、TLI和RMSEA只能识别组内模型的误设，对模式系数（pattern coefficients）更灵敏，而SRMR-W则对因子协方差（factor covariance）误设更灵敏。而且在组内因子协方差误设问题上，TLI表现好于CFI和RMSEA。Hsu（2009）建议使用SRMR-W结合RMSEA和CFI评估组内模型。另外，SRMR-B是目前唯一对组间模型误设灵敏的拟合指数，而且对因子协方差比模式系数的误设更灵敏（Hsu et al，2015）。不过，在ICC过小的情况下，SRMR-B对组间模型误设识别敏感性降低，也不推荐使用（Hsu，2009；Boulton，2011；Hsu et al，2016）。

同时，也可以借助输出结果中卡方值、LL值进行模型比较，从而评估模型相对的改进情况。M*plus* 中，MLM和MSEM估计方法基本通用。基于卡方值或LL进行模型比较的内容，参见第6章第10节 。

8.2　多水平因子分析

8.2.1　多水平因子分析概述

因子分析是应用广泛的多变量（multivariate）统计技术，可以用于探索数据背后的模式、验证假设、数据降维等方面。因子分析可分为探索性因子分析（exploratory factor analysis）和验证性因子分析（confirmatory factor analysis）。从形式上讲，如果多因子

模型中，对特定因子载荷进行约束，如等于0或彼此相等，那么就是约束的或者验证性因子分析模型。

对于存在聚类效应的数据，单水平因子分析忽略嵌套结构。由于组群内部个体的相似性，传统因子分析因为观测值独立性假设被破坏，卡方检验统计量、因子载荷、标准误有偏，统计显著性检验失效，因子得分也会有偏。而且这种偏差随着ICC增加而增大（Julian 2001；Dyer et al，2005；Preacher et al，2010）。而且，因为嵌套数据破坏独立性假定，即使去除潜变量影响，量表题项也仍然相关，这会影响研究者对量表是否单维的判断（Dyer et al，2005）。根据模拟研究（Julian，2001）的结论，一般ICC大于0.05且组群规模相对较大时，应该考虑将因子分析纳入多水平框架。

按照Muthén（1989；1990）的定义方式，多水平验证性因子分析可以表达为：

个体层：

$$Y_{ij} = \nu_j + \boldsymbol{\Lambda}_W \boldsymbol{\eta}_{ij} + \boldsymbol{\varepsilon}_W \qquad (8.23)$$

组群层：

$$\boldsymbol{\upsilon}_j = \boldsymbol{\nu} + \boldsymbol{\Lambda}_B \boldsymbol{\eta}_j + \boldsymbol{\varepsilon}_B \qquad (8.24)$$

复合模型：

$$Y_{ij} = \boldsymbol{\nu} + \boldsymbol{\Lambda}_B \boldsymbol{\eta}_j + \boldsymbol{\Lambda}_W \boldsymbol{\eta}_{ij} + \boldsymbol{\varepsilon}_B + \boldsymbol{\varepsilon}_W \qquad (8.25)$$

其中，Y_{ij} 代表第 j 个组群第 i 个个体的响应，v 是总均值，Λ_W 为组内因子载荷矩阵，Λ_B 为组间因子载荷矩阵，η_{ij} 是组内随机变动的因子，η_j 是组间随机变动的因子，ε_W 和 ε_B 分别为组内和组间的独特因素。[(12)]

以 Y_B 代表组均值，Y_W 代表个体与所在组群均值的偏差，那么每个个体的观测值可以分解为两部分：

$$Y_T = Y_B + Y_W \quad (8.26)$$

依循Cronbach（1976）的思路，响应向量 Y_{ij} 的总体总协方差矩阵 Σ_T，也可以分解为总体组间协方差矩阵 Σ_B 和总体组内协方差矩阵 Σ_W，即：

$$\Sigma_T = \Sigma_B + \Sigma_W \quad (8.27)$$

分不同水平来看，即：

$$\boldsymbol{\Sigma}_W = \boldsymbol{\Lambda}_W \boldsymbol{\Psi}_W \boldsymbol{\Lambda}'_W + \boldsymbol{\Theta}_W \quad (8.28)$$

$$\boldsymbol{\Sigma}_B = \boldsymbol{\Lambda}_B \boldsymbol{\Psi}_B \boldsymbol{\Lambda}'_B + \boldsymbol{\Theta}_B \quad (8.29)$$

其中，$\boldsymbol{\Psi}_W$ 和 $\boldsymbol{\Psi}_B$ 分别代表组内和组间的因子协方差矩阵。Λ_W 和 Λ_B 分别是组内和组间因子载荷协方差矩阵。Θ_W 和 Θ_B 分别是组内和组间残差协方差对角矩阵。

为了得到模型参数，需要对 Σ_B 和 Σ_W 进行估计。逻辑上，可以将样本数据也分解为组间变异和组内变异两个部分，其协方差结构为：

$$S_T = S_B + S_W \tag{8.30}$$

其中，

$$S_T = \frac{\sum_j^J \sum_i^{n_j} (Y_{ij} - \bar{Y})(Y_{ij} - \bar{Y})'}{N - 1} \tag{8.31}$$

$$S_B = \frac{\sum_j^J n_j(\bar{Y}_j - \bar{Y})(\bar{Y}_j - \bar{Y})'}{J - 1} \tag{8.32}$$

$$S_W = \frac{\sum_j^J \sum_i^{n_j} (Y_{ij} - \bar{Y}_j)(Y_{ij} - \bar{Y}_j)'}{N - J} \tag{8.33}$$

Muthén（1989；1990）证明，不能简单地用S_B 和S_W 作为Σ_B 和Σ_W 的估计。假定在J个平衡组群中，每个组群规模均为n，总样本量即为$N=nJ$ 。定义两个样本协方差矩阵：S_{PW} ，混合组内协方差矩阵（pooled within covariance matrix）；$\mathbf{S}_B^*$ ，调整的组间协方差矩阵（scaled between covariance matrix）。

总体组内协方差矩阵Σ_W 的无偏估计是S_{PW} ，可通过如下公式得到：

$$S_{PW} = \frac{\sum_j^J \sum_i^n (Y_{ij} - \bar{Y}_j)(Y_{ij} - \bar{Y}_j)'}{N - J} \tag{8.34}$$

调整的组间协方差矩阵S_B^* ，可通过如下公式得到：

$$S_B^* = \frac{\sum_j^J n(\bar{Y}_j - \bar{Y})(\bar{Y}_j - \bar{Y})'}{J - 1} \tag{8.35}$$

可以证明，S_{PW} 是Σ_W 的最大似然估计值，样本量为$N-J$；S_B^* 是$\Sigma_W + c\Sigma_B$ 的最大似然估计值，样本量为J，c 等于组群规模n 。即：

$$S_{PW} = \hat{\Sigma}_W \tag{8.36}$$

$$S_B^* = \hat{\Sigma}_W + c\hat{\Sigma}_B \tag{8.37}$$

有 $\hat{\Sigma}_B = c^{-1}(S_B^* - S_{Pw})$ (8.38)

由于不同的组群样本量不同，导致模型过于复杂。Muthén（1989；1990）建议忽略非平衡，采取一个变通的方式，令：

$$c^* = \frac{N^2 - \sum_j^J n_j^2}{N(J - 1)} \tag{8.39}$$

单水平因子分析模型（single-level factor analysis）基于总体协方差矩阵（S_T）。Muthén（1994）证明，总体协方差矩阵是组内和组间因子结构的混合，当使用总体协方差矩阵进行因子分析时，组群水平因子结构以及因子载荷估计有偏。

忽略嵌套结构，基于样本组间协方差矩阵（between-group covariance matrix，S_B），采取聚合方式，在组群水平进行因子分析，如果进行自上而下的推论，与对嵌套数据使用聚合回归来解释个

体层面问题一样，存在“层次谬误”。而且，从式8.32或式8.35来看，利用组均值的因子分析正是基于加权或未加权的组均值样本协方差矩阵，S_B 或S_B^* 正是用变量聚合所得组均值（并中心化）得到的。由于样本组间协方差矩阵（S_B ）并非总体组间协方差矩阵（Σ_B ）的无偏估计，其不仅是总体组间协方差矩阵的函数，同时也是加权的组群规模以及组内协方差矩阵（Σ_W ）的函数（见式8.37）（Muthen，1994）。也就是说包含了来自组内和组间的方差成分。此时如果采用聚合方式形成组群水平构念，假定组内协方差矩阵（Σ_W ）是误差或噪声方差（noise variance）的函数，结果会低估组间结构，因子载荷产出现保守性偏误（Dyer et al，2005）。按8.1.2节所述，这样的处理属于MMC方式，在组内样本数较少、ICC较小或者属于反映性（reflective）构念测量等情况下，存在缺乏信度等问题。

8.2.2　多水平验证性因子分析

由于多水平因子分析模型比较复杂，经常迭代难以收敛，Muthén（1989；1994）建议采取五步法构建多水平因子分析模型：

● 基于总体协方差矩阵（S_T）执行传统的CFA。由于存在嵌套数据结构，这一步的参数估计和拟合统计量可能有偏，尤其是存在较大组内相关、组规模较大、变量高度相关、组内因子结构和组间因子结构不同的情况下。

● 组间变异估计。这一步要回答“多水平分析是否适合数据”的问题。可通过计算ICC来判断。如果所有题项的ICC都接近于0，就不必

使用MCFA。这一步中，通常将组间协方差矩阵和组内协方差矩阵存为新的数据，供后续步骤使用。

● 基于样本混合组内协方差矩阵（S_{PW}）进行因子分析。S_{PW} 通过减去所在组的均值排除了组间差异，只反映组内因子结构。如果总变异主要来源于组间，那么使用S_{PW} 的拟合可能不如基于S_T 的因子分析。反之，如果总变异主要来源于组内，那么使用S_{PW} 的拟合相较基于S_T 的因子分析可能会有所改善。

● 基于样本组间协方差矩阵（S_B）进行因子分析。这一步提供了有关组群水平因子结构的信息。实际中，Σ_B 估计经常非正定，可通过对S_B 的估计反映Σ_B 的结构。

● 多水平验证性因子分析。基于上述步骤结果，进行多水平验证性因子分析，同时对组内和组间进行分析。

不考虑理论意义和指标的可聚合性，以下我们用PISA数据中学生学习策略（ST27Q01—13）作为多水平验证性因子分析示例。量表共13个题项，简单起见，这里选取其中的8个题项，见表8-4。量表测度1-4级，近似视为连续变量，数据已经将负向计分调整为正向计分。经计算，总体KMO＝0.8941，适合做因子分析。

表8-4　学习策略量表题项

变量名	标　签	均值	标准差	ICC
ST27Q01	记住文本中的每个东西	2.505	0.899	0.037
ST27Q03	尽可能多地记住细节	2.872	0.902	0.030
ST27Q04	将信息与以往的知识建立关联	2.542	0.946	0.011
ST27Q05	读课文很多次直到能复述	1.944	0.918	0.056
ST27Q07	一遍又一遍地读文本	2.472	0.976	0.030
ST27Q08	搞清楚信息如何在课外派上用场	1.963	0.916	0.028
ST27Q10	将材料与自身经验建立联系以便更好理解	2.254	0.963	0.009
ST27Q12	搞清楚信息如何与现实生活相适应	2.202	0.930	0.015

各题项中，ST27Q01、ST27Q03、ST27Q05、ST27Q07与记忆（memorisation ）有关；ST27Q04、ST27Q08、ST27Q10、ST27Q12与理解（elaboration）有关。这里针对水平-1，设定为双因子结构。一般来讲，组群水平的因子往往呈现出聚合的现象，简化起见可采用单因子结构。其实际意义即：在学校层面，该量表视作对学生学习策略的整体测量。对于多水平因子分析而言，一个重要的优势是根据理论，可以对不同水平设定差异性的因子结构，而不是强加等结构的假定。这在某些研究中，具有重要意义。

图8-3即根据上述假设制作的组内双因子、组间单因子的两水平验证性因子分析模型。组内模型由因子指向指标的箭头头部的圆点代表随机截距，对应于各指标在组间模型中相应的潜变量。

为了更好地理解上述说法，这里需要提醒的是，不要为M*plus* 代码中BY命令的形式所迷惑。虽然BY命令左侧是因子，右侧是指标，但实际上从测量模型来看，潜变量决定指标，路径图上看也就是代表因果关系的箭头是从潜变量指向指标。所以，实际上BY命令所表达的是某因子与相应指标之间的若干方程。表示因子载荷的箭头代表的是斜率。

在多水平随机截距模型中，组内方程的截距是随机变化的，在组间不同，也就是表现在水平-2的层次，方差存在于水平-2。因此，实际上，从MSEM的建模思路上来看，这些随机截距其实就是相应水平-1指标的组间效应。它们可以作为组间因子的指标，受其作用。

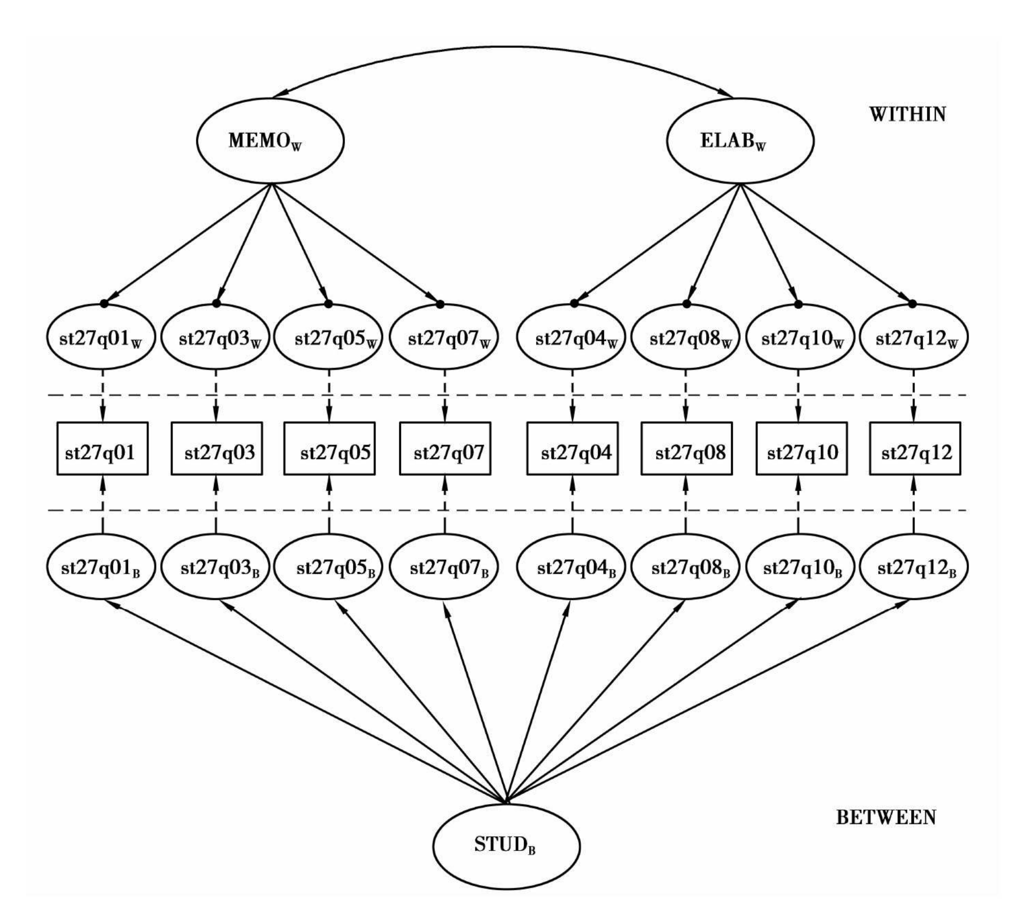


图8-3　两水平验证性因子分析

第一步进行一个单水平的因子分析，M*plus* 代码见框8-3。

框8-3　MCFA的第一步

```
TITLE：MCFA STEP1，Confirmatory factor analysis sample;
DATA:   FILE IS pisafactor.dat;
```

```
VARIABLE:
  Names are
          schoolid  st27q01  st27q03  st27q04  st27q05  st27q07
st27q08 st27q10 st27q12
      hisei girl sector;
  USEVARIABLES＝st27q01-st27q12;
  Missing are all（－9999）;
ANALYSIS: ESTIMATOR＝ML;
MODEL:
  MEMO BY st27q01 st27q03 st27q05 st27q07;
  ELAB BY st27q04 st27q08 st27q10 st27q12;
OUTPUT: SAMPSTAT STDYX MODINDICES RESIDUAL;
```

基于量表实际意义和探索性因子分析结果考虑，在个体层面，各题项测量了两个潜变量，可以区分出两个维度。在学校层面，可以考虑作为对学校学生学习策略整体性的测量，即一个因子。

第二步，存储组内和组间协方差矩阵，M*plus* 代码见框8-4。

框8-4 MCFA的第二步

```
TITLE: MCFA STEP2, Create within and between matrices and
obtain ICC values;
DATA:  FILE IS pisafactor.dat;
VARIABLE:
  Names are
```

```
            schoolid  st27q01  st27q03  st27q04  st27q05  st27q07
st27q08 st27q10 st27q12
        hisei girl sector;
  USEVARIABLES=st27q01 st27q12;
  Missing are all (-9999);
  CLUSTER is schoolid;
ANALYSIS: TYPE=TWOLEVEL BASIC;
          ESTIMATOR=ML;
OUTPUT: SAMPSTAT;
SAVEDATA: SIGBETWEEN IS BetCov.dat;   ! 保存组间协方差矩阵
          SAMPLE    IS WinCov.dat;   ! 保存组内协方差矩阵
```

代码中的SIGBETWEEN和SAMPLE是SAVEDATA命令的选项。其中SAMPLE在TYPE＝TWOLEVEL的情况下，存储组内协方差Σ_W矩阵；SIGBETWEEN存储组间协方差Σ_B矩阵。[13]执行上述代码后，工作路径中新增了存储组内和组间协方差矩阵的数据文件。整理后结果见表8-5和表8-6。

表8-5　组内协方差矩阵

	ST27Q01	ST27Q03	ST27Q04	ST27Q05	ST27Q07	ST27Q08	ST27Q10	ST27Q12
ST27Q01	0.7776							
ST27Q03	0.4402	0.7898						
ST27Q04	0.2597	0.3117	0.8835					
ST27Q05	0.2945	0.2841	0.2544	0.7970				
ST27Q07	0.2947	0.3402	0.2238	0.4478	0.9237			
ST27Q08	0.1641	0.1439	0.3622	0.2449	0.2076	0.8157		
ST27Q10	0.1861	0.1770	0.4433	0.2284	0.2107	0.4539	0.9184	
ST27Q12	0.1745	0.1796	0.4074	0.2110	0.1970	0.4787	0.5431	0.8508

表8-6 组间协方差矩阵

	ST27Q01	ST27Q03	ST27Q04	ST27Q05	ST27Q07	ST27Q08	ST27Q10	ST27Q12
ST27Q01	0.0296							
ST27Q03	0.0254	0.0240						
ST27Q04	0.0016	0.0050	0.0102					
ST27Q05	0.0311	0.0275	0.0022	0.0470				
ST27Q07	0.0245	0.0234	0.0046	0.0305	0.0283			
ST27Q08	0.0103	0.0116	0.0060	0.025	0.0124	0.0232		
ST27Q10	0.0056	0.0066	0.0054	0.0130	0.0062	0.0124	0.0083	
ST27Q12	0.0060	0.0079	0.005	0.0156	0.0072	0.0169	0.0088	0.0133

续表

根据第二步模型，多数题项有一定程度的组内相关，不过个别题项很低（见表8-4）。可以理解，学习策略因为主要是反映个人特质，量表各题项聚类效应不是很突出。

第三步，根据组内协方差矩阵进行组内因子分析，见框8-5。

框8-5 MCFA的第三步

```
TITLE： MCFA STEP3， Procedural factor estimated with within matrix;
DATA: FILE IS WinCov.dat;              ！读取组内协方差矩阵
    TYPE IS COVA;    ！定义分析的矩阵类型
    NOBSERVATIONS IS 4492;      ！组内样本量，N－G，本例中为4644－152
VARIABLE： NAMES ＝ st27q01 st27q03 st27q04 st27q05 st27q07 st27q08 st27q10 st27q12;
ANALYSIS: ESTIMATOR＝ML;
```

```
MODEL:
    MEMOw BY st27q01 st27q03 st27q05 st27q07;
    ELABw BY st27q04 st27q08 st27q10 st27q12;
OUTPUT: SAMPSTAT STDYX MODINDICES RESIDUAL;
```

组内模型的估计代码中，MEMOw和ELABw作为水平-1的因子，默认是相关的，因为它们属于外生变量。同时，默认设置各因子分别与各自第1个指标之间的载荷路径系数为1。因子的方差默认估计。

第三步模型反映的是组内特征。因子负荷略有差别，不过从某些指标来看，拟合相较于第一步模型并没有明显改善。

第四步，根据组间协方差矩阵进行组间因子分析，见框8-6。

框8-6 MCFA的第四步

```
TITLE: MCFA STEP4, Procedural factor estimated with between-
cluster matrix;
DATA: FILE IS BetCov.dat;              ! 读取组间协方差矩阵
    TYPE IS COVA;      ! 定义分析的矩阵类型
    NOBSERVATIONS IS  152;      ! 组间样本量G，本例中为152
VARIABLE : NAMES = st27q01  st27q03  st27q04  st27q05  st27q07
st27q08 st27q10 st27q12;
ANALYSIS: ESTIMATOR=ML;
MODEL:
     STUDb BY st27q01 st27q03 st27q04 st27q05 st27q07 st27q08
```

```
st27q10 st27q12;
OUTPUT: SAMPSTAT STDYX MODINDICES RESIDUAL;
```

第四步模型为组间模型。拟合情况较差，甚至出现Loglikelihood值异常、残差协方差矩阵非正定的情况，可能与变量ICC过小、组间协方差矩阵元素值较小，以及组群层面样本量较少有关。

本例中单因子（x^2 ＝1652.376，df ＝20）和双因子（x^2 ＝1059.149，df ＝19）拟合结果存在差异。这里根据假设，组间仍采取单因子结构。

第五步，根据原始数据进行多水平因子分析，见框8-7。

框8-7　MCFA的第五步

```
TITLE: Multilevel confirmatory factor analysis;
DATA: FILE IS pisafactor.dat;
VARIABLE:
  Names are
          schoolid  st27q01  st27q03  st27q04  st27q05  st27q07
st27q08 st27q10 st27q12
      hisei girl sector;
    USEVARIABLES＝st27q01 st27q12;
    Missing are all（－9999）;
    CLUSTER is schoolid;
    WITHIN＝;
```

```
    BETWEEN＝;
ANALYSIS：TYPE＝TWOLEVEL;
      ESTIMATOR＝ML;
MODEL：%WITHIN%
      MEMOw BY st27q01 st27q03 st27q05 st27q07;
      ELABw BY st27q04 st27q08 st27q10 st27q12;
    %BETWEEN%
         STUDb  BY  st27q01  st27q03  st27q04  st27q05  st27q07
st27q08 st27q10 st27q12;
OUTPUT：SAMPSTAT STDYX MODINDICES RESIDUAL;
```

第五步模型为使用联立方法估计的最终的多水平验证性因子分析，使用的是原始数据。注意，这里代码的语法符合MLC模型的规范。由于ST27Q01-ST27Q12各指标需要分解到不同水平，所以变量声明部分未指定属于哪一层面。WITHIN和BETWEEN分别构建相应测量模型（用BY表示），指标名称不加区分，但实际上分别对应组内和组间不同部分。另外要注意的是，组内和组间的潜变量名称上要区别开，表示不同水平的构念。以上各步骤的相关输出结果进行整理后得到表8-7。

表8-7　多水平验证性因子分析结果

	步骤 1		步骤 3		步骤 4	步骤 5		
	Total		Within		Between	Within		Between
	F1	F2	$F1_W$	$F2_W$	F_B	$F1_W$	$F2_W$	F_B
ST27Q01	0.676*** (0.012)		0.664*** (0.011)		0.373*** (0.069)	0.670*** (0.013)		0.850*** (0.082)
ST27Q03	0.692*** (0.012)		0.682*** (0.011)		0.463*** (0.063)	0.688*** (0.012)		0.896*** (0.086)
ST27Q05	0.651*** (0.012)		0.637*** (0.012)		0.758*** (0.034)	0.632*** (0.013)		0.995*** (0.059)
ST27Q07	0.642*** (0.012)		0.632*** (0.012)		0.470*** (0.063)	0.630*** (0.013)		0.913*** (0.061)
ST27Q04		0.630*** (0.011)		0.634*** (0.011)	0.365*** (0.070)		0.635*** (0.011)	0.283 (0.213)
ST27Q08		0.709*** (0.009)		0.702*** (0.009)	1.007*** (0.004)		0.705*** (0.009)	0.938*** (0.085)
ST27Q10		0.766*** (0.008)		0.765*** (0.008)	0.886*** (0.018)		0.766*** (0.008)	0.868*** (0.204)
ST27Q12		0.790*** (0.008)		0.787*** (0.008)	0.955*** (0.008)		0.789*** (0.008)	0.894*** (0.137)
Corr	0.499*** (0.015)		0.503*** (0.015)			0.492*** (0.016)		
Model Fit								
LL	−44472.725		−42700.807		1104.616	−44381.879		
AIC	88995.450		85435.615		−2177.233	88845.757		
BIC	89156.533		85544.586		−2128.850	89109.934		
CFI	0.920		0.918		0.351	0.917		
RMSEA	0.104		0.104		0.733	0.074		
SRMR	0.049		0.055		0.261			
SRMR-Within						0.056		
SRMR-Between						0.223		
N	4644		4644		4644	4644		

标准化系数；括号内为标准误

+ $p<0.1$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

从策略上来讲，上述步骤实际上结合了分离法（segregating approaches）和联立法。上述各步估计方法均采用最大似然法（ML）。M*plus* 中对于利用组间协方差矩阵和组内协方差矩阵（术语

称为summary data）进行的估计（即第3、4步），只能使用ML、GLS和ULS方法，其他估计方法需要原始数据。

从相关指标来看，相对于单水平CFA，MCFA模型拟合较好。各因子负荷基本上也在0.5以上且显著。比较第3、4、5步所得结果可以看到，第3步组内因子分析结果与第5步组内部分结果十分接近，但第4步组间因子分析部分结果与第5步组间部分结果有较大差异。虽然由于考虑ICC、样本量、模型复杂性等多种因素后，两种方法究竟哪一个更有效存在争议，但一般来讲结果应该比较接近。根据相关研究（Yuan＆Bentler，2007；Schweig，2014），当模型较简单、ICC很低，同时组内样本量较小时，分离法相对更有效。不过本例中，组间和组内样本规模均属于较高水平。

从拟合结果来看，多水平验证性因子分析（步骤5）整体优于不考虑嵌套结构的单水平因子分析（步骤1）。对于结果的评估，由于可以分别输出不同层面模型的拟合指标，分离法在这方面有优势。从第3、4步输出的指标来看，组内模型拟合尚可，但组间模型拟合情况较差。根据第5模型拟合结果中的SRMR-Between可以得到同样结论。另外，比较步骤3、4和5相应模型输出结果可以看到，组内部分结果接近，但组间部分结果有较大差异。在既定模型设定下，原因应该主要在于数据自身。就本例而言，如果作为实际应用的话，需要考虑各指标ICC大小和组群层面变量之间的结构关系，重新评估对此数据使用多水平因子分析的必要性。从理论意义上来讲，实际上多水平因子分析更适用于对参照点位于水平－2的构念，也就是具有共享性特征变量（Kozlowski＆Klein，2000；Klein＆Kozlowski，2000）、反映性变量（Lüdtke et al，2008）或者氛围变量（Marsh et al，2012）进行

分析。这类测量一般ICC较大，且组间模型也更有实质意义。显然本例中的构念参照点典型位于个体层面。

另外正如前面所指出的，实际上从Muthén（1990；1994）有关组间和组内协方差矩阵的定义式来看，尽管组间协方差矩阵估计经过校正，但本质上作为组间构念的各指标还是通过对观测指标聚合而成（同时进行了中心化处理），所以Marsh et al（2009）称之为潜-显模型（latent-manifest model），也就是虽然是基于多指标的潜测量构建，但水平-2构念是通过显聚合（manifest aggregation）而成。相比之下，基于MLC的思路，水平-2构念的构建使用的是潜聚合（latent aggregation），方法上更具优势。不过MLC方法属于单指标测量，对于带有多指标的MCFA或者一般MSEM的测量模型部分，各指标分解后，在测量模型中相对于因子而言处于响应变量的位置，已不属于协变量。“双潜多水平模型”概念避免了上述问题，更准确、更一般性地刻画出MSEM建模思路的关键特征。

8.2.3　多水平探索性因子分析

当对因子结构缺乏明确的事前假定或可供参考的相关文献时，可以考虑使用多水平探索性因子分析。

当数据存在嵌套结构时，单水平的探索性因子分析因为不能考虑到数据中存在的相关性，可能会导致对题项之间的相关或协方差高估，并导致因子得分有偏。而且，即使多水平因子分析，组群层因子结构和个体层因子结构也未必一致。因子分析需要注意到这种复杂性。

Mplus 支持变量为连续变量、二分变量、有序或无序类别变量、截取变量、计数变量等多种类型的因子分析，不过具体估计方法需要根据具体情况进行选择。对于两水平探索性因子分析，如果是截取、无序类别或计数变量，目前尚不支持。[(14)]

同样基于上述学生学习策略量表，进行多水平探索性因子分析。*Mplus* 代码见框8-8，其中“ANALYSIS”的“TYPE”部分，“TWOLEVEL EFA 1 2 UW1 2 UB”代表的意思是，进行两水平探索性因子分析，组内、组间因子数均设定为3种情况：1个因子、2个因子和不约束。这样，可能的组合情况就是2^3，也就是8种类型。

框8-8　多水平探索性因子分析

```
TITLE: MEFA for PISA;
DATA:   FILE IS pisafactor.dat;
VARIABLE:
  Names are
          schoolid st27q01 st27q03 st27q04 st27q05 st27q07
st27q08 st27q10 st27q12
      hisei girl sector;
  USEVARIABLES＝st27q01 st27q12;
  Missing are all（－9999）;
  CLUSTER＝schoolid;
ANALYSIS:
    TYPE＝TWOLEVEL EFA 1 2 UW 1 2 UB;   ！数字代表组间模型因子数
        ！UW代表组内模型给出非约束类型
```

```
        ！UB代表组间模型给出非约束类型
    ESTIMATOR＝ML;
    ROTATION＝QUARTIMIN;          ！四方最小法
                                    ！多水平EFA中不能使用VARIMAX和
PROMAX
                                  ！默认GEOMIN
```

组内双因子-组间单因子的组合相应输出结果见框8-9。

框8-9　多水平探索性因子分析输出结果（组内2因子，组间1因子）

```
EXPLORATORY FACTOR ANALYSIS WITH 2 WITHIN FACTOR(S) AND 1 BETWEEN FACTOR(S):

WITHIN LEVEL RESULTS

           QUARTIMIN ROTATED LOADINGS (* significant at 5% level)
                  1            2
              ________     ________
ST27Q01         0.714*      -0.022
ST27Q03         0.770*      -0.045*
ST27Q04         0.255*       0.504*
ST27Q05         0.500*       0.154*
ST27Q07         0.540*       0.075*
ST27Q08        -0.001        0.711*
ST27Q10        -0.008        0.770*
ST27Q12        -0.036*       0.816*


BETWEEN LEVEL RESULTS

           QUARTIMIN ROTATED LOADINGS (* significant at 5% level)
                  1
              ________
ST27Q01         0.901*
ST27Q03         0.945*
ST27Q04         0.074
ST27Q05         0.997*
ST27Q07         0.937*
ST27Q08         0.916*
ST27Q10         0.823*
ST27Q12         0.878*
```

综合各种组合输出模型的拟合指标来看，如果不计UNRESTRICTED的情况，相对而言，组内-组间双因子的组合拟合情况属于相对较好的类型，组内双因子-组间单因子的组合次之。具体结果详见表8-8。

表8-8　读书态度量表多水平探索性因子分析结果

Type	Within	Between	LL	AIC	CFI	RMSEA	SRMR-W	SRMR-B
1	1	1	−45523.884	91127.768	0.727	0.133	0.114	0.276
2	2	1	−44212.080	88518.160	0.945	0.066	0.039	0.251
3	UNR	1	−43886.229	87820.457	0.998	0.000	0.000	0.256
4	1	2	−45491.189	91076.378	0.732	0.145	0.114	0.162
5	2	2	−44195.296	88498.593	0.947	0.073	0.039	0.126
6	UNR	2	−43869.680	87801.361	1.000	0.000	0.000	0.086
7	1	UNR	−45493.757	91035.515	0.730	0.111	0.114	0.000
8	2	UNR	−44187.932	88437.864	0.947	0.051	0.039	0.000

8.2.4　多水平因子分析拓展

以下为几种常见的更复杂形式的多水平因子分析模型相应M*plus*代码，包括：带有随机斜率的多水平验证性因子分析、带协变量的多水平因子分析（MIMIC）和多水平多组因子分析。

（1）带有随机斜率的多水平验证性因子分析

以上MCFA模型示例均属于随机截距模型，实际上，M*plus* 中可以允许MCFA因子载荷随机变动。框8-10是带有随机斜率的MCFA示例。

框8-10　带有随机斜率的多水平验证性因子分析

```
TITLE: Multilevel CFA with Varying Loadings;
DATA:   FILE IS pisafactor.dat;
VARIABLE:
  Names are
```

```
          schoolid  st27q01  st27q03  st27q04  st27q05  st27q07
st27q08 st27q10 st27q12
      hisei girl sector;
  USEVARIABLES＝st27q01 st27q12;
  Missing are all（－9999）;
  CLUSTER is schoolid;
  WITHIN＝;
  BETWEEN＝;
ANALYSIS: TYPE＝TWOLEVEL RANDOM;
      ESTIMATOR＝BAYES;
      ALGORITHM＝INTEGRATION;
      PROCESSORS＝4;
      FBITERATIONS＝（1000）;
MODEL: %WITHIN%
      s1-s4|MEMOw BY st27q01 st27q03 st27q05 st27q07;
      MEMOw@1;
      s5-s8|ELABw BY st27q04 st27q08 st27q10 st27q12;
      ELABw@1;
      %BETWEEN%
          STUDb  BY  st27q01  st27q03  st27q04  st27q05  st27q07
st27q08 st27q10 st27q12;
      STUDb@1;
OUTPUT: TECH1 TECH8;
```

在一般的测量模型中，默认设置第1个因子的载荷路径系数为1。上述多水平验证性因子分析带有随机斜率，为了令模型可识别，不能

再采取固定路径系数法，只能采用固定方差法，所以出现了“MEMOw@1”等代码。另外，由于存在随机斜率，模型复杂，这里采取的是贝叶斯估计方法。

（2）带协变量的多水平因子分析

在原MCFA基础上，水平-1增加学生性别（girl）协变量，水平-2增加学校类型（sector）协变量。两个变量均属于所在层面专用变量，所以需要在VARIABLE部分指定所属层次。相应代码见框8-11。

框8-11　带协变量的多水平因子分析（MIMIC）

```
TITLE: MCFA with covariates
DATA: FILE IS pisafactor.dat;
VARIABLE:
  Names are
          schoolid  st27q01  st27q03  st27q04  st27q05  st27q07
st27q08 st27q10 st27q12
      hisei girl sector;
  USEVARIABLES＝st27q01 st27q12 girl sector;
  Missing are all（－9999）;
  CLUSTER is schoolid;
  WITHIN＝girl;
  BETWEEN＝sector;
ANALYSIS: TYPE＝TWOLEVEL;
```

```
      ESTIMATOR＝ML；
MODEL：
    %WITHIN%
      MEMOw BY st27q01 st27q03 st27q05 st27q07；
      ELABw BY st27q04 st27q08 st27q10 st27q12；
      MEMOw ELABw ON girl；
    %BETWEEN%
          STUDb  BY  st27q01  st27q03  st27q04  st27q05  st27q07
st27q08 st27q10 st27q12；
      STUDb ON sector；
OUTPUT：TECH4  STDYX；
```

（3）多水平多组因子分析

多水平多组（multi-group）因子分析，常用来分析等值型或不变性（invariance）问题。一个需要注意的问题是，在实际操作时，数据处理需要区别分组变量来自组内还是组间。

对于分组变量属组间水平，不同组群之间所有系数均独立的情况，方程式如下（Asparouhov＆Muthén，2012）。

$$Y_{ijg} = Y_{Wijg} + Y_{Bjg} \tag{8.40}$$

$$Y_{Wijg} \sim N(\mu_{1g}, \Sigma_{1g}) \tag{8.41}$$

$$Y_{Bjg} \sim N(\mu_{2g}, \Sigma_{2g}) \tag{8.42}$$

$$Y_{Wijg} = \mu_{1g} + \Lambda_{wg}\eta_{wijg} + \varepsilon_{ijg} \tag{8.43}$$

$$Y_{Bjg} = \mu_{2g} + \Lambda_{bg}\eta_{bjg} + \varepsilon_{jg} \tag{8.44}$$

似然函数也是群各自似然函数加总。

$$LL = \sum_{g=1}^{G} LL_g \tag{8.45}$$

对于分组变量属组内水平，较为复杂，式8.44不成立，方程8.42变为（Asparouhov＆Muthén，2012）：

$$Y_{Bj.} \sim \mathrm{N}(\mu_{2.}, \Sigma_{2.}) \tag{8.46}$$

或

$$Y_{Bjg} \sim \mathrm{N}(\mu_{2g}, \Sigma_{2g}) \tag{8.47}$$

$$\mathrm{Cov}(Y_{Bj1}, Y_{Bj2}) = \rho \tag{8.48}$$

因此，两种情况需要区别对待。框8-12是分组变量在组间的多水平多组因子分析示例。

框8-12　多水平多组因子分析：分组变量在组间

```
TITLE: Multilevel Multi-group CFA;
DATA:   FILE IS pisafactor.dat;
VARIABLE:
```

```
  Names are
          schoolid st27q01 st27q03 st27q04 st27q05 st27q07
st27q08 st27q10 st27q12
      hisei girl sector;
  USEVARIABLES＝st27q01 st27q12 sector;
  Missing are all（－9999）;
  WITHIN＝;
  BETWEEN＝;
  GROUPING＝sector（0＝public 1＝private）;
  CLUSTER＝schoolid;
ANALYSIS:
  TYPE＝TWOLEVEL;
  ESTIMATOR＝ML;
MODEL:
  %WITHIN%
    MEMOw BY st27q01 st27q03 st27q05 st27q07;
    ELABw BY st27q04 st27q08 st27q10 st27q12;
  %BETWEEN%
     STUDb BY st27q01 st27q03 st27q04 st27q05 st27q07 st27q08
st27q10 st27q12;
OUTPUT: TECH4 STDYX;
```

在多组因子分析模型中，默认的是测量不变性，因子载荷和截距被设置为跨组相等。如果需要释放该假设，需另外设定组特定（group-specified）模型。参见“基础篇”（王孟成，2014：183-185）。

框8-13是分组变量在组内的多水平多组因子分析示例。其中，groupid的需要事先生成，如使用Stata进行预处理，命令为：egen groupid＝group（girl schoolid），实际上就是以性别（0、1编码）和学校编号进行交叉分类。

框8-13　多水平多组因子分析：分组变量在组内

```
TITLE：Multilevel Multi-group CFA;
DATA:   FILE IS pisafactor2.dat;
VARIABLE:
  Names are
     schoolid st27q01 st27q03 st27q04 st27q05 st27q07 st27q08
st27q10 st27q12
    hisei girl sector groupid;
  USEVARIABLES＝st27q01 st27q12 girl groupid;
  Missing are all（－9999）;
  WITHIN＝;
  BETWEEN＝;
  GROUPING＝girl（0＝boy 1＝girl）;
  CLUSTER＝groupid;      ！必须针对组内分组变量对组群重新编码，否则报
错Cluster ID
          ！cannot appear in more than one group.
ANALYSIS: TYPE＝TWOLEVEL;
      ESTIMATOR＝ML;
MODEL:
    %WITHIN%
```

```
    MEMOw BY st27q01 st27q03 st27q05 st27q07;
    ELABw BY st27q04 st27q08 st27q10 st27q12;
  %BETWEEN%
      STUDb BY st27q01 st27q03 st27q04 st27q05 st27q07
st27q08 st27q10 st27q12;
OUTPUT: TECH4 STDYX;
```

8.3　多水平路径分析

8.3.1　单水平路径分析

有时变量之间的因果关系复杂，一个变量对于某些变量可能是原因变量，而对于另外一些变量是结果变量，这时不能简单地以因变量或自变量来划分变量类型，对于这样的模型，可以用路径分析方法来解决。路径分析便于研究一些观测变量之间的复杂的因果关系，用来估计并检验变量之间的直接与间接效应。

路径分析可以看作结构方程的一个特例。当假定不存在测量误差、潜变量时，结构方程即为路径分析模型。比如，我们要研究PISA2009美国数据中学生的阅读得分问题，数据中阅读得分有5个似真值。这里选取其均值作为最终结果变量，关注阅读兴趣（joyread，实为st24q01 st24q11综合而成，ICC＝0.03）、性别（girl）的影响。同时，模型中也包括学校层面变量学生性别比（pcgirls）。基于上述假定，构建如下路径图（见图8-4）。

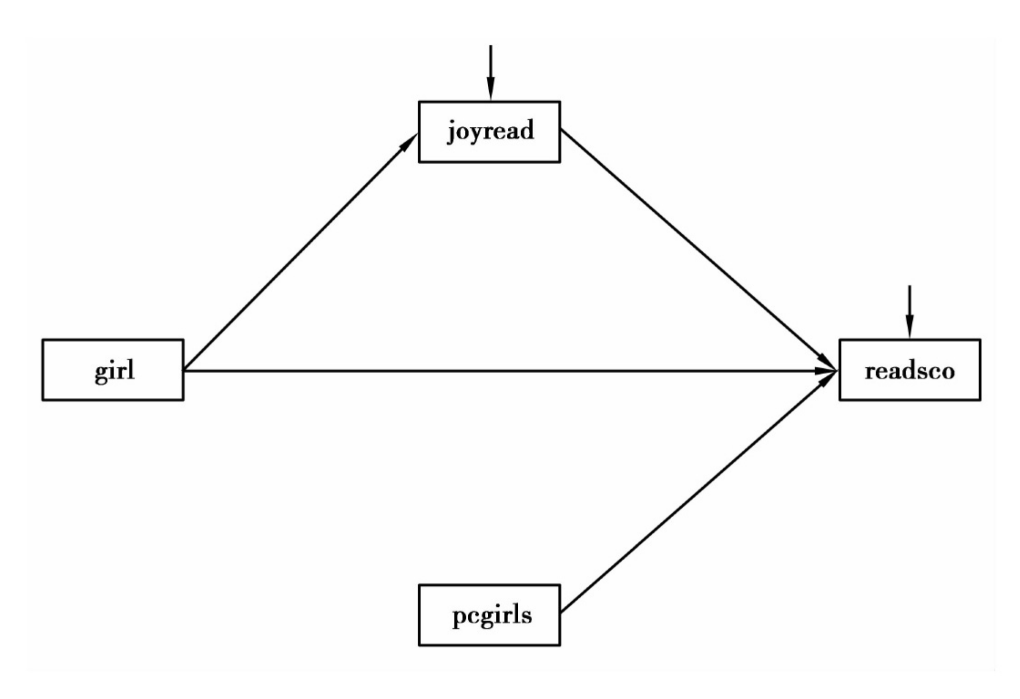


图8-4　阅读得分的路径分析模型

按照一般路径分析模型对待，相应M*plus* 代码见框8-14。

框8-14　路径分析

```
DATA:   File is pisampa.dat;
Variable:
  Names are schoolid joyread pcgirls girl readsco;
  Missing are all (－9999) ;
  USEVARIABLES＝joyread pcgirls girl readsco;
MODEL:
    readsco ON joyread pcgirls girl;
    joyread ON girl pcgirls;
    girl WITH pcgirls@0;
Model indirect:
```

```
  readsco ind joyread girl;
OUTPUT: SAMPSTAT STDYX;
```

单水平路径分析可以分解为若干回归方程，对于递归模型，一般有几个内生变量就对应几个方程。根据假设，上述模型有readsco、joyread两个内生变量，所以需要设定两个方程。ON左手边代表内生变量，右手边则是相应的预测变量。按照M*plus* 语法规范，默认估计外生变量之间的协方差，这里假定外生变量girl与pcgirls之间不相关，所以需要设定其间协方差为0。[15]

从输出结果（Output结果省略，具体可参见表8-8）来看，假设得到验证，模型拟合良好。对间接效应的估计也初步可以判断成立。然而对于这个具有嵌套结构的数据，如果存在不容忽视的聚类效应，可能会导致错误的统计推论。另外，单水平路径分析把个体与学校的变量混在一起，没有区分出变量的层次。

8.3.2　多水平路径分析

多水平路径分析，整合了多水平分析和路径分析，考虑到了数据嵌套结构框架下变量之间复杂的因果关系。在M*plus* 中，多水平路径分析允许随机截距和随机斜率。

典型的多水平路径分析，原则上应该是只有内生变量有组间变化的模型（对应相应的随机截距）。水平-1观测变量很难在水平-2没有对应的潜变量，除非相应ICC几乎为0，或者依照假设设定变量无组间变化即只属于组内专用变量，因此如果使用MLC策略引入潜协变量，模

型可视为一般意义上的MSEM。不过，一旦设置joyread有随机截距，意味着其也是组内/组间两用变量，组间也存在方差。这样，即使组间不设置相应子模型，M*plus* 默认也将在结果的组间部分输出两个内生变量的协方差（READSCO WITH JOYREAD）。这里在组间设置了阅读得分与阅读兴趣的方程，从而可以估计后者对前者的组间效应。所以，实际的代码中VARIABLE定义部分，只是规定了学生性别比（pcgirls）作为组间专用变量，学生性别（girl）作为组内专用变量。由此，重新在多水平框架下构建阅读得分的路径分析模型，路径图见图8-5。

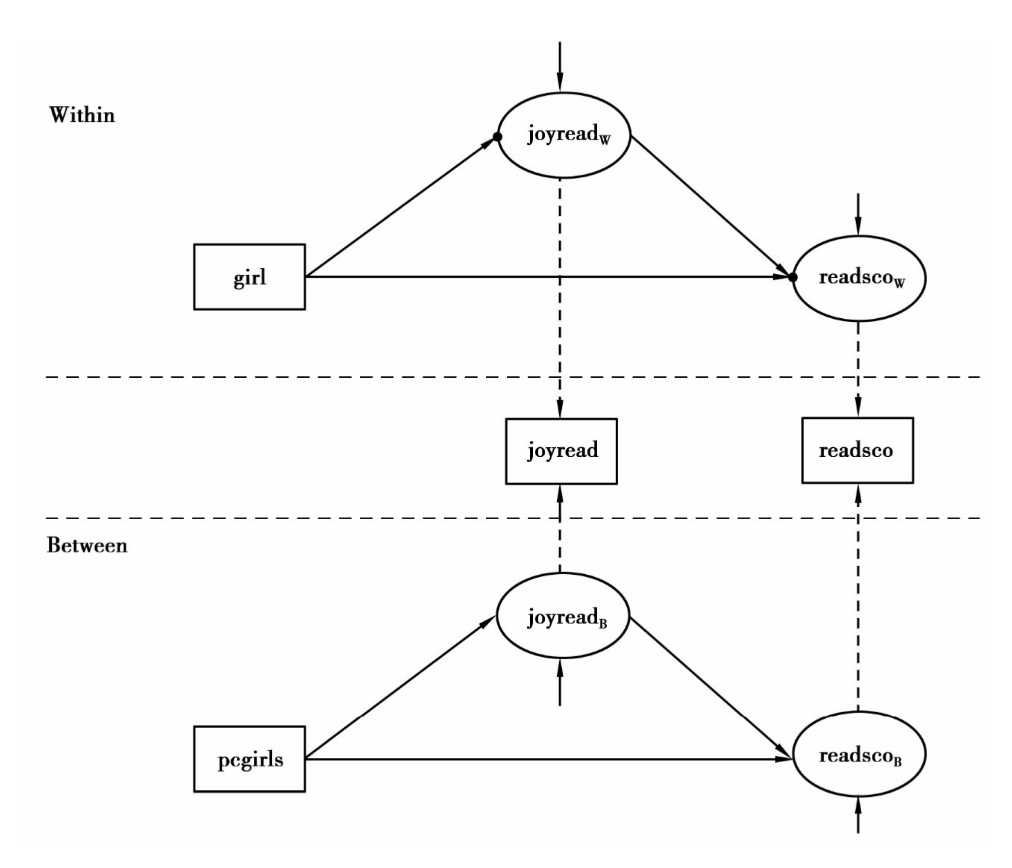


图8-5　阅读得分的多水平路径分析

简化起见，构建随机截距多水平路径分析模型。相应M*plus* 代码见框8-15。

框8-15　多水平路径分析

```
TITLE: Multilevel Path Analysis with Random Intercept;
DATA:   File is pisampa.dat;
Variable:
  Names are schoolid joyread pcgirls girl readsco;
  Missing are all (－9999) ;
  USEVARIABLES＝joyread pcgirls girl readsco;
  WITHIN＝girl;
  BETWEEN＝pcgirls;
  CLUSTER＝schoolid;
ANALYSIS:
  TYPE＝TWOLEVEL;
  ESTIMATOR＝ML;
MODEL:
  %WITHIN%
    readsco ON joyread girl;
    joyread ON girl;
  %BETWEEN%
    joyread readsco ON pcgirls;
    readsco ON joyread;
OUTPUT: SAMPSTAT STDYX;
```

注意，上述代码变量界定部分（VARIABLE）中，除了最终结果变量（即readsco），也没有指定joyread所属的层次，因此属于组内/组间两用的变量，分析时会自动分解到组间和组内两部分。组内模型中，readsco和joyread变量上的圆点代表相应方程截距有随机的变

化。二者的随机截距实际上存在于各自的组间部分。在模型组间部分，二者跨组变动，或者说方差体现于水平-2。组间模型描述了它们如何受到学校层面因素的影响，二者在组间也存在结构关系。

注意，这里pcgirls作为水平-2整体变量，刻画的是学校某方面的整体特征。非由学生性别（girl）观测值聚合而成或分解出组间成分得到，其在组内没有方差变化，因此不能同时出现在MODEL部分的WITHIN和BETWEEN层面，只能在水平-2参与模型构建。同理，如果水平-1变量没有聚类效应，ICC接近于0，可设定为仅在组内变动的变量（Only-W）。这里性别即做了这样的处理。

输出结果（部分）见框8-16。整理为回归表格（见表8-8），其中包括了单水平路径分析结果。

框8-16　多水平路径分析输出结果（部分）

MODEL RESULTS				
				Two-Tailed
	Estimate	S.E.	Est./S.E.	P-Value
Within Level				
READSCO ON				
JOYREAD	34.249	1.097	31.217	0.000
GIRL	6.661	2.285	2.915	0.004
JOYREAD ON				
GIRL	0.660	0.030	21.851	0.000
Residual Variances				
JOYREAD	0.972	0.021	46.087	0.000
READSCO	4964.642	107.805	46.052	0.000
Between Level				
JOYREAD ON				
PCGIRLS	0.003	0.003	0.949	0.343
READSCO ON				
PCGIRLS	−0.782	0.470	−1.663	0.096
JOYREAD	117.670	27.917	4.215	0.000
Intercepts				
JOYREAD	−0.477	0.138	−3.456	0.001
READSCO	566.076	26.128	21.666	0.000
Residual Variances				
JOYREAD	0.033	0.008	4.330	0.000
READSCO	1287.344	201.334	6.394	0.000

表8-9　学生阅读得分多水平路径分析

	单水平	多水平	
		组内	组间
Readsco			
Intercepts	531.725*** (7.778)		566.076*** (26.128)
Joyread	37.122*** (1.187)	34.249*** (1.097)	117.670*** (27.917)
Girl	2.931 (2.526)	6.661** (2.285)	
Pcgirls	−0.602*** (0.162)		−0.782+ (0.470)

续表

	单水平	多水平	
		组内	组间
Joyread			
Intercepts	−0.474*** (0.096)		−0.477** (0.138)
Girl	0.652*** (0.030)	0.660*** (0.030)	
Pcgirls	0.003 (0.002)		0.003 (0.003)
Model Fit			
LL	−51470.147	−31280.951	
AIC	102966.295	62585.902	
BIC	103049.843	62662.545	
CFI	0.937	1.000	
RMSEA	0.142	0.000	
SRMR	0.041		
SRMR-Within		0.000	
SRMR-Between		0.000	

括号内为标准误

+ $p<0.1$，* $p<0.05$，** $p<0.01$，*** $p<0.001$

从结果来看，在单水平路径分析模型中，在个体层面，阅读兴趣影响阅读得分在统计上显著，但性别对阅读得分影响不显著，不过通过阅读兴趣对阅读成绩的间接效应显著。[16]多水平路径分析结果中，阅读兴趣与阅读得分之间的关系在两个层面均显著。同时在水平-1，不同性别个体阅读得分存在显著差异。

对比单水平和多水平路径分析结果可以看到，阅读兴趣对阅读成绩之间的关系，在单水平路径分析结果中，该系数为37.122。而从多水平路径分析结果来看，实际上该值混淆了组内和组间效应。从数量来看，组间效应（117.670）高于组内效应（34.249）。

根据模型相关拟合指标（见表8-8）输出结果，多水平路径分析模型拟合要优于单水平路径分析模型。

对于多水平模型，M*plus* 均不输出路径图。如果得到上述路径分析结果，可以在概念图基础上使用相关制图软件进行加工，将路径系数等信息标注清楚。

8.4　多水平结构方程模型

8.4.1　单水平结构方程模型

结构方程模型可以像路径分析那样考虑复杂的因果关系，同时纳入潜变量，容许测量误差。

同样是上面有关学生阅读得分的问题，使用阅读得分似真值（pv1read—pv5read）。纳入水平-1性别（girl）、社会经济地位

（hisei）、阅读态度（简化起见，去掉了3个负荷相对较小的指标，保留st24q01、st24q02、st24q03、st24q05、st24q06、st24q07、st24q08、st24q11），以及基于水平-2测量的教师行为（sc17q01、sc17q03、sc17q05、sc17q06、sc17q09、sc17q11、sc17q13）和学生行为（sc17q02、sc17q04、sc17q07、sc17q08、sc17q10、sc17q12）（负向计分）。

作为参照，首先构建单水平结构方程模型（见图8-6），M*plus* 相应代码见框8-17。

框8-17　阅读得分的单水平结构方程模型

```
TITLE: Structural Equation Model;
DATA: File is pisamsem.dat;
Variable:
  Names are schoolid st24q01 st24q02 st24q03 st24q05 st24q06
st24q07 st24q08 st24q11
     hisei pv1read pv2read pv3read pv4read pv5read sc17q01
sc17q02 sc17q03
     sc17q04 sc17q05 sc17q06 sc17q07 sc17q08 sc17q09 sc17q10
sc17q11 sc17q12
    sc17q13;
  Missing are all（－9999）;
    USEVARIABLES ＝ pv1read  pv5read  st24q01  st24q02  st24q03
st24q05 st24q06 st24q07
        st24q08 st24q11 hisei sc17q01 sc17q13;
```

```
ANALYSIS:
  ESTIMATOR=MLR;
  ITERATIONS=30000;
ODEL:
  ReadSco BY pv1read pv5read;
  ReadAtt BY st24q01 st24q02 st24q03 st24q05 st24q06 st24q07
st24q08 st24q11;
  TeacBeh BY sc17q01 sc17q03 sc17q05 sc17q06 sc17q09 sc17q11
sc17q13;
   StudBeh  BY  sc17q02  sc17q04  sc17q07  sc17q08  sc17q10
sc17q12;
  ReadSco ON ReadAtt TeacBeh StudBeh hisei;
  ReadAtt ON TeacBeh StudBeh hisei;
  StudBeh ON TeacBeh hisei;
  TeacBeh WITH hisei;
OUTPUT: STDYX;
```

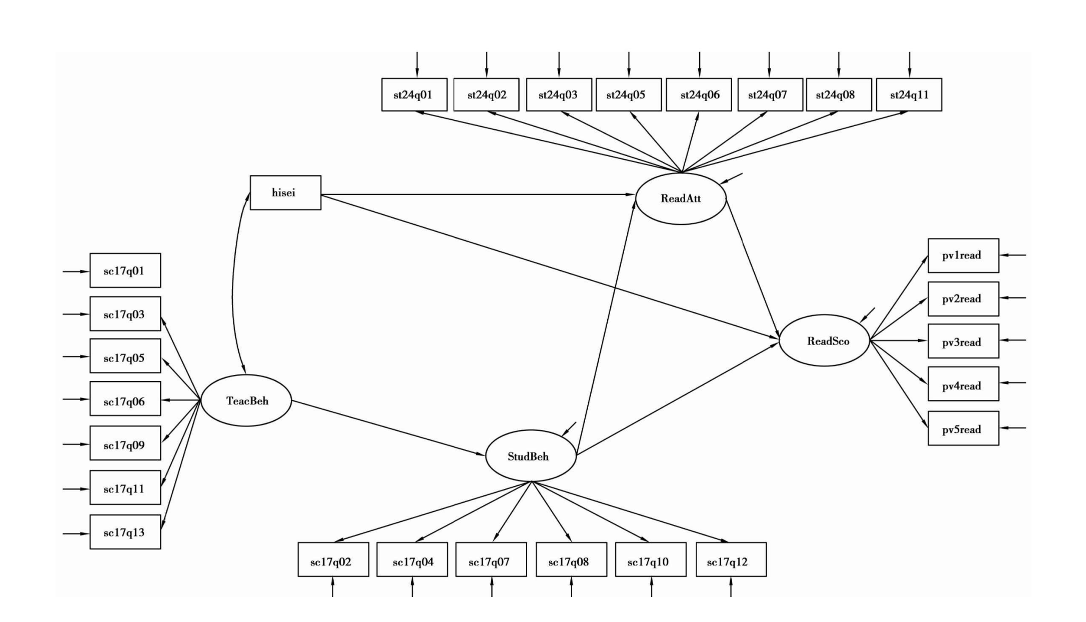


图8-6　阅读得分的结构方程模型

这是一个单水平结构方程模型，MODEL部分，BY语句构建测量模型，ON语句构建潜变量之间的结构方程。

从输出结果（见表8-7）来看，模型拟合尚好。从各路径系数假设检验情况判断，基本上可以支持假设。不过，由于该数据存在嵌套结构，单水平结构方程模型实际上可能存在问题。

8.4.2　多水平结构方程模型

根据假定，我们构建如下多水平结构方程模型（见图8-5）。与单水平结构方程模型相比，多水平结构方程模型分为组间和组内两个层面。

与单水平结构方程模型不同，这里学校风气、性别比作为学校特征专属于水平-2，基于个体测量的学生家庭社会经济地位、阅读策略属于组内/组间两用变量，性别作为水平-1专用变量。

相应的M*plus* 代码见框8-18。

框8-18　阅读得分的多水平结构方程模型

```
TITLE : Multilevel Structural Equation Model with Random
Intercept;
DATA: File is pisamsem.dat;
Variable:
```

```
  Names are schoolid st24q01 st24q02 st24q03 st24q05 st24q06
st24q07 st24q08 st24q11
     hisei pv1read pv2read pv3read pv4read pv5read sc17q01
sc17q02 sc17q03
     sc17q04 sc17q05 sc17q06 sc17q07 sc17q08 sc17q09 sc17q10
sc17q11 sc17q12
    sc17q13;
Missing are all (－9999);
USEVARIABLES＝pv1read-pv5read st24q01 st24q02 st24q03 st24q05
st24q06 st24q07
        st24q08 st24q11 hisei sc17q01-sc17q13;
  WITHIN＝;
  BETWEEN＝sc17q01-sc17q13;
  CLUSTER＝schoolid;
ANALYSIS:
  TYPE＝TWOLEVEL;
  ESTIMATOR＝MLR;
  MCONVERGENCE＝0.1;
MODEL:
  %WITHIN%
  ReadScow BY pv1read-pv5read;
  ReadAttw BY st24q01 st24q02 st24q03 st24q05 st24q06 st24q07
st24q08 st24q11;
  ReadScow ON ReadAttw hisei;
  ReadAttw ON hisei;
%BETWEEN%
  ReadScob BY pv1read-pv5read;
```

```
  ReadAttb BY st24q01 st24q02 st24q03 st24q05 st24q06 st24q07
st24q08 st24q11;
  TeacBeh BY sc17q01 sc17q03 sc17q05 sc17q06 sc17q09 sc17q11
sc17q13;
  StudBeh BY sc17q02 sc17q04 sc17q07 sc17q08 sc17q10
sc17q12;
  ReadScob ON ReadAttb TeacBeh StudBeh hisei;
  ReadAttb ON TeacBeh StudBeh hisei;
  StudBeh ON TeacBeh hisei;
  TeacBeh WITH hisei;
OUTPUT: STDYX;
```

多水平结构方程模型的M*plus* 命令中，要区分WITHIN和BETWEEN不同层面。像多水平因子分析一样，不同水平会构建对应子模型。

如果需要在两个水平同时考虑阅读态度和阅读得分，也就是分别评价学生个体以及学校层面学生整体的阅读态度和阅读得分水平，需要在不同的水平构建相应的测量方程。hisei以及pv1read—pv5read、st24q01—st24q11各指标在两个水平均出现，属于组内／组间两用的类型，在变量界定部分不能指定所属层次。这体现了MLC的思路，也就是对观测协变量进行多水平的分解。在不同水平的方程中，这些变量虽以相同的变量名出现，但分别表示各自的组内和组间部分。不过由于这些指标在特定水平对应不同的潜变量，而这些构念分属不同层次，所以潜变量在名称上要注意进行区分。比如本例中，ReadAttw代表组内阅读兴趣，ReadAttb代表组间阅读兴趣。教师行为和学生行为属于水平-2的构念，是在水平-2层面收集的，也就是只是对学校某方

面整体状况的度量，在组内无方差变化，所以设定为只属于BETWEEN部分。

M*plus* 中，MSEM默认使用MLR进行估计，输出稳健 χ^2 统计量和稳健标准误。在中等程度违反分布假定的情况下，可以提供无偏的估计。如果因变量为类别变量，ML需要高维数字积分，可能会出现迭代不收敛、估计不准确的问题，此时可以采用WLSM（V）法进行估计。M*plus* 高版本也支持MSEM中随机斜率的设定，需要设定数字积分算法。部分输出结果见框8-19。

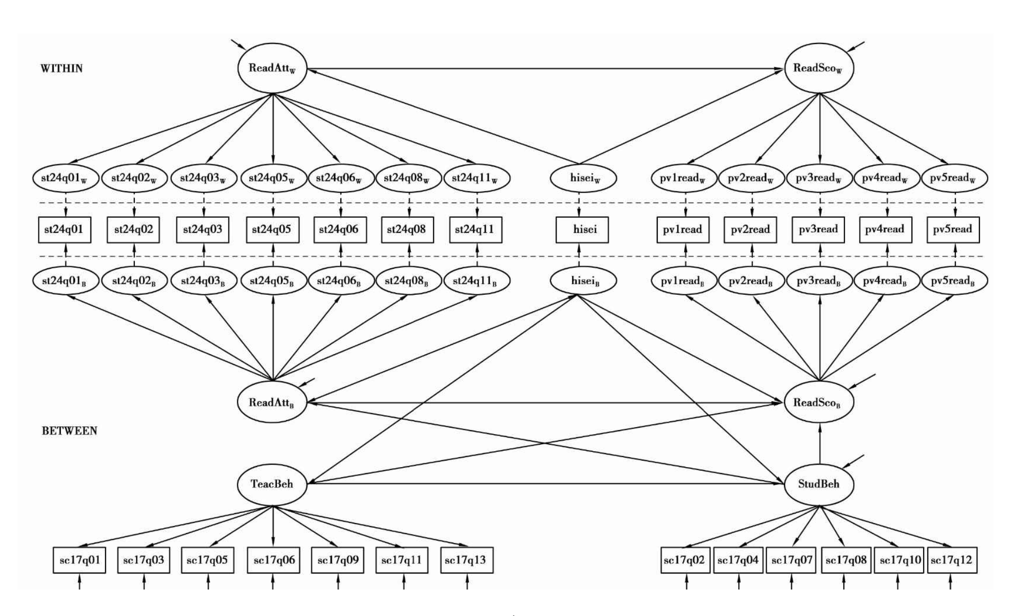


图8-7　阅读得分的多水平结构方程模型

框8-19　多水平结构方程模型输出结果（部分）

```
STANDARDIZED MODEL RESULTS


STDYX Standardization

                                                         Two-Tailed
                    Estimate       S.E.      Est./S.E.    P-Value
Within Level
 READSCOW    ON
 READATTW              0.463      0.014       33.897      0.000
 READSCOW    ON
   HISEI               0.169      0.015       11.362      0.000
 READATTW    ON
   HISEI               0.107      0.017        6.432      0.000
 Variances
   HISEI               1.000      0.000      999.000    999.000
Between Level
 READSCOB    ON
   READATTB            0.511      0.324        1.577      0.115
   TEACBEH            -0.110      0.163       -0.677      0.498
   STUDBEH             0.155      0.172        0.902      0.367
 READATTB    ON
   TEACBEH             0.219      0.274        0.800      0.424
   STUDBEH            -0.325      0.285       -1.140      0.254
 STUDBEH    ON
   TEACBEH             0.659      0.066        9.920      0.000
 READSCOB    ON
   HISEI               0.461      0.328        1.406      0.160
 READATTB    ON
   HISEI               0.810      0.165        4.896      0.000
 STUDBEH    ON
   HISEI               0.288      0.076        3.800      0.000
 TEACBEH    WITH
   HISEI               0.355      0.087        4.087      0.000
 Means
   HISEI               8.246      0.604       13.654      0.000
```

整理结果见表8-10。从输出的拟合指数来看，MSEM拟合情况要优于单水平SEM。不过组间模型拟合相对较差。

表8-10 PISA学生阅读得分结构方程

	单水平	多水平	
		组内	组间
Readsco			
Readatt	49.501 *** (1.602)	48.236 *** (1.626)	262.679 (298.813)
hisei	1.410 *** (0.075)	0.910 *** (0.078)	3.141 (2.302)
Teacbeh	−4.961 (5.148)		−9.209 (13.367)
Studbeh	30.066 *** (4.972)		12.393 (13.808)
Readatt			
hisei	0.006 *** (0.001)	0.006 *** (0.001)	0.011 * (0.005)
Teacbeh	−0.001 (0.049)		0.036 (0.044)
Studbeh	−0.021 (0.048)		−0.051 (0.057)
Studbeh			
hisei	0.003 *** (0.000)		0.024 ** (0.008)
Studcli			
Teaccli	0.802 *** (0.026)		0.688 *** (0.120)
Teaccli with hisei	1.119 *** (0.127)		1.123 ** (0.336)
Model Fit			
LL	−217101.289	−172642.595	
AIC	434382.577	345525.190	
BIC	434959.782	346294.797	
CFI	0.944	0.984	
RMSEA	0.036	0.024	
SRMR	0.036		
SRMR-Within		0.024	
SRMR-Between		0.152	

非标准化系数，省略测量方程、截距；括号内为标准误

+ $p<0.1$，* $p<0.05$，** $p<0.01$，*** $p<0.001$

对比单水平和多水平SEM结果会发现，单水平SEM与多水平SEM中某些变量的显著性有所不同。比如，在组内方程中，家庭社会经济地位对学生阅读得分作用显著。但在组间方程中相应系数不显著，不过后者的效应值大概还是前者的3.5倍。具体来看，$p=0.172$，如果学校样本量增加，该系数应该会统计上显著。社会经济地位与阅读兴趣之间，组内和组间的系数均统计上显著，组间系数大概是组内系数的2倍。另外还可以看到，在单水平SEM中，学生行为（StudBeh）对阅读得分的作用是显著的。然而，从MSEM的角度来看，原则上该变量作为水平-2变量，直接作用于学生个体阅读得分是有问题的，混淆了组间和组内效应。实际上，StudBeh只有组间方差，也就是说只能在水平-2变化。对某个固定的学校而言其实是一个定值，因此对水平-1变量并不能产生直接影响。从实际结果来看，在MSEM中，该变量作用于结果变量的组间效应，系数并不显著。

较新版本的M*plus* 也可以允许构建路径系数带有随机斜率的MSEM（Muthén＆Asparouhov，2008；Preacher et al，2010）。比如，增加组内hisei对ReadAttw的随机斜率，可以在“ReadAttw ON hisei”语句前面加上“s | ”，同时再将ANALYSIS部分设置成“TYPE＝TWOLEVEL RANDOM；”。不过一般来讲，带有随机效应的MSEM较为复杂，容易造成迭代不收敛等问题。下一节将在较简单的情形下考虑纳入带有随机斜率的MSEM。

8.5　本章小结

MSEM模型结合了MLM和MSEM的优势，引入了潜变量以及多水平潜协变量的概念，同时考虑到多水平回归模型进行情景分析时存在的测量

误差和抽样误差，并可分析更复杂的因果关系，因此在方法上更具优势。

本章首先介绍了MSEM的基本原理、估计方法、拟合评价指标等问题。并结合实例，对作为特例的多水平验证性因子分析、多水平探索性因子分析、多水平路径分析和一般的MSEM建模进行了示范。

由于MSEM在多水平分析中具备的相对优势，可以预见这类研究将会超越经典的MLM。一些实际应用案例和相关最新进展，参见本章“推荐阅读”部分提到的相关文献。

※推荐阅读※

多水平因子分析属于多水平结构方程的特例。这方面早期文献可参见Muthén（1991）。研究实例如Dyer et al（2005）、Reise et al（2005）、Gajewski et al（2010）、Dunn et al（2015）、Barbour＆Lammers（2015）、Davidov et al（2016）等。Kim et al（2016）示范了MCFA结果汇报的格式，并对从1994至2014之间72个使用MCFA的经验研究进行了评述。

基于多水平结构，以等值型或不变性为入手点，MCFA／MSEM常用于跨文化或群体差异比较研究，参见Cheung＆Au（2005）、Cheung et al（2006）、Selig et al（2008）、Feskens＆Hox（2011）、Davidov et al（2013）的介绍。

多水平结构方程模型在某些方面已经显现出超越多水平回归模型的优势，这方面的基础文献参见Goldstein＆Mcdonald（1988）、

Rabe-Hesketh et al（2004）、Mehta＆Neale（2005）、Muthén＆Asparouhov（2008）、Toit et al（2008）、Lüdtke et al（2008）、Muthén＆Asparouhov（2011）、Stapleton（2013）等。晚近的一些MSEM的研究在方法方面不断深化。如Depaoli & Clifton（2015）、Holtmann et al（2016）探讨了MSEM估计的贝叶斯方法，Hsu et al（2016）评价ICC对于水平特定拟合指标（基于partially saturated model method）功效的影响，Geldhof et al（2014）介绍了针对嵌套数据利用多水平验证性因子分析估计组特定信度（level-specific reliability）的指标和方法，等等。Marsh et al（2014）以有关班风的教育研究为例，Stapleton et al（2016）使用有关经验和模拟数据，探讨了构念分属于构成性／共享型、或情境／氛围等不同类型情况下进行多水平建模的思路，对于相关模型的选择与设定具有指导意义。

应用MSEM的经验研究正逐渐增多，如Franzini et al（2005）、Walsh＆Matthews（2010）、Uthman et al（2010）、Beauchamp et al（2014）、Huhtala et al（2015）、Wardell et al（2015）、Zhang et al（2016）、Ratanavaraha et al（2016）等。

---

(1)　注意，因为需要允许截距随机变化，多水平结构方程模型属于带有均值结构的全模型。

(2)　在基本的意义上，显变量（manifest variables）即直接观测所得变量，有时也称为指标（indicators）、题项（item）、观测变量（observed variables）等。比如，身高、年龄、性别等。潜变量（latent variables）即非直接观测的变量，有时也称为非观测变量（unobserved variable）、构念（construct）、因子（factor）等。比如，智商、福利、数学能力等。潜变量代表了基础的、深层的构念，或者指标之间的关系。在更广泛意义上，误差（error）、未观测异质性（unobserved heterogeneity）、缺失值（missing

data）、反事实结果（counterfactuals）等，也属于潜变量范畴。按照Skrondal＆Rabe-Hesketh（2004，1）的界定，人们无法在现实当中观测的随机变量均属于潜变量。

(3) 在MSEM框架中，会使用与随机效应ANOVA一致的方法（Muthén＆Asparouhov，2011），将变量自动分解为组内／组间两个独立、正交的部分。其中，组间部分即β_{0j}，用组间元素表示即$\gamma_{00}+\mu_{0j}$，组内部分即ε_{ij}（见式6.21、式6.22）。当然，实际估计中需要分解观测变量的方差（Pituch＆Stapleton，2011；Stapleton，2013），也就是将总方差分解为组间方差$Var(\mu_{ij})=\tau_{00}$和组内方差$Var(\varepsilon_{ij})=\sigma^2$两个部分。

(4) 该式也可以表达为$X_{ijw}=X_{ij}-X_{jb}$，可看作对组内潜协变量进行潜组均值的中心化（Muthén＆Muthén，2012：269）。

(5) 水平-2变量由于没有组内（W）部分，只作为组间效应（Only-B）变量。当然也存在只有组内效应的变量，但这种情况假定观测值独立，较为少见（Preacher et al，2016）。组间／组内专用变量也可以视作相应组内／组间部分变差为0。

(6) MSEM的框架下允许结果变量可以处于水平-2，这一点多水平回归模型也无法做到。

(7) 按照Mehta＆Neale（2005）所讲，FIML优势的关键是其似然值以个体数据向量为基础进行计算，由此为设定和估计更复杂的模型奠定了条件。而且，如果将组群作为观测单位，组群内的个体作为变量（各组个体不等，相当于非平衡数据），多水平回归模型也可以表达为一个结构方程模型，适合使用FIML进行拟合。Bauer（2003）提到将MLM作为SEM进行估计的做法（类似潜增长模型）。Christian et al（2013）则使用MSEM估计潜增长模型。

(8) 不过一般来讲，对量表进行多水平验证性因子分析，随机斜率应该不显著，否则意味着量表缺乏跨组群的等值性（Hox，2010：299）。

(9) 考虑到变量测量层次等因素具体方法选择参见M*plus* 手册（Muthén＆Muthén，2012：601-602）。

(10) 注意，目前M*plus* 中，带有随机斜率的MSEM不输出TLI、RMSR等拟合指标。具体分析中可使用相应固定斜率模型进行估计。实际上，单就代码来说，M*plus* 中并不区分MSEM和MLM，所以没有随机斜率的MLM也会有TLI、RMSR等拟合指标输出，但因为MLM只有显变量且往往是饱和模型，所以输出的这类指标一般也没有意义。

(11)　实际操作中可以尝试利用M*plus* 的TYPE＝TWOLEVEL BASIC方式，通过设定SAVEDATA：SIGBETWEEN / SAMPLE参数，分别保存组间协方差矩阵和组内协方差矩阵（具体操作详见下一节多水平验证性因子分析部分），然后在此基础上分别对水平-1和水平-2建模，并根据相关输出结果评价不同水平模型的拟合情况（见本章第2节 多水平验证性因子分析结果）。类似上述所提的第2种方案。

(12)　以上为随机截距模型，新的估计方法允许因子载荷Λ_W也可以在不同组群之间变动。

(13)　注意M*plus* 中不同方法估计所得矩阵有所不同，参见M*plus* 手册（Muthén＆Muthén，2012：749）。

(14)　具体参见M*plus* 手册（Muthén＆Muthén，2012：601-602）。

(15)　关于单水平路径分析模型，M*plus* 代码书写规范参见王孟成（2014：33-37）。

(16)　不过这里需要提醒的是，在多水平模型情境下，中介效应的检验要复杂得多，规范的分析方法详见本书第9章 。

9　多水平中介效应分析

9.6　本章小结

9.1　多水平中介效应分析概述

科学研究的主要任务就是检验变量之间的因果关系。在确立因果关系问题时，分析自变量到因变量之间的机制问题尤为重要。有关如何厘清因果机制的问题，统计上即对应着中介效应分析。考虑X 对Y 的影响，如果X 通过对M 的影响对Y 起作用，则M 即中介变量。比如，家庭社会地位通过影响学生阅读兴趣间接影响学生的阅读成绩。

一般的中介效应分析仅针对单水平数据而言，在多水平数据的情况下，水平-1中X 与Y 之间的系数是水平-2变量的函数，往往随机变动。忽视数据的多水平结构和相似性将导致效应估计有偏，标准误低估，从而第Ⅰ类错误概率增加。在多水平模型中，中介效应的两种表示$c-c'$ 和ab 并不相等（Krull＆MacKinnon，1999）。

以两水平中介模型为例，根据X、Y 和M 所在的水平不同，理论上说可能的中介模型有8个类型（Krull＆MacKinnon，2001；Preacher et al，2010；温忠麟等，2012：230），如表9-1所示。

表9-1　多水平中介效应的类型

类　型	X	M	Y	MLM 局限
2-2-2	水平-2	水平-2	水平-2	没必要
2-2-1	水平-2	水平-2	水平-1	两步法
2-1-2	水平-2	水平-1	水平-2	不可用
2-1-1	水平-2	水平-1	水平-1	混淆
1-2-2	水平-1	水平-2	水平-2	不可用
1-2-1	水平-1	水平-2	水平-1	不可用
1-1-2	水平-1	水平-1	水平-2	不可用
1-1-1	水平-1	水平-1	水平 -1	混淆

一般来讲，实际中很少出现较低水平变量影响较高水平变量的情况。[(1)]因此，具体应用中常见的主要是2-2-2、2-2-1、2-1-1、1-1-1共4个类型。而对于2-2-2类型，可以采取聚合方式在水平-2进行分析。所以研究中常见多水平中介效应模型的类型主要是2-2-1、2-1-1、1-1-1这3个类型（Zhang et al，2009）。相应模型图见图9-1至图9-3。

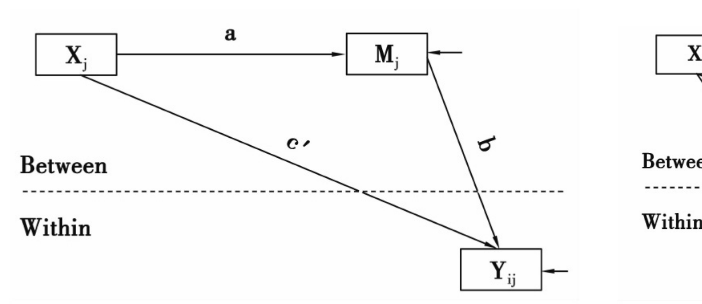


图9-1　多水平中介2-2-1模型

图9-2　多水平中介2-1-1模型

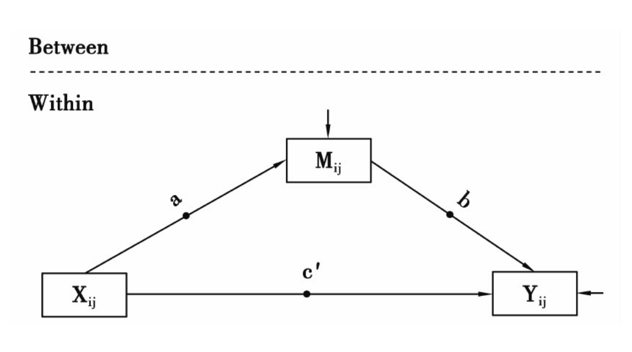


图9-3　多水平中介1-1-1模型

其中，多水平2-2-1中介模型属于高水平中介模型，2-1-1和1-1-1属于低水平中介模型。区分的依据即中介变量所处的层次。也正是由于中介变量的位置，决定了前者不含随机路径系数，后者含随机路径系数（Bauer et al，2006）。进一步，多水平2-2-1随机中介效应模型属于固定效应模型，因其只含随机截距。

分析多水平中介效应，可以采取MLM和MSEM两种方式。与MLM方式相比，使用MSEM进行中介效应分析，具有可以纳入对误差的考虑、不会混淆组内与组间中介效应、可以允许结果变量位于水平-2等优势（Preacher et al，2010；2011）。

9.2　基于MLM的多水平中介效应分析

9.2.1　多水平固定中介效应分析

常见多水平固定中介效应模型（Krull＆MacKinnon，2001；Zhang et al，2009）包括以下几种。

2-2-1模型：[2]

$$Y_{ij} = \beta_{0j}^{(1)} + \varepsilon_{ij}^{(1)}$$

$$\begin{aligned} \beta_{0j}^{(1)} &= \gamma_{00}^{(1)} + \gamma_{01}^{(1)} X_j + \mu_{0j}^{(1)} \\ M_j &= \beta_0^{(2)} + \gamma_{01}^{(2)} X_j + \varepsilon_j^{(2)} \\ Y_{ij} &= \beta_{0j}^{(3)} + \varepsilon_{ij}^{(3)} \\ \beta_{0j}^{(3)} &= \gamma_{00}^{(3)} + \gamma_{01}^{(3)} X_j + \gamma_{02}^{(3)} M_j + \mu_{0j}^{(3)} \end{aligned} \tag{9.1}$$

就此模型而言，c 即$\gamma_{01}^{(1)}$；c' 即$\gamma_{01}^{(3)}$；a 即$\gamma_{01}^{(2)}$；b 即$\gamma_{02}^{(3)}$。

2-1-1模型：

$$
\begin{aligned}
Y_{ij} &= \beta_{0j}^{(1)} + \varepsilon_{ij}^{(1)} \\
\beta_{0j}^{(1)} &= \gamma_{00}^{(1)} + \gamma_{01}^{(1)} X_j + \mu_{0j}^{(1)} \\
M_{ij} &= \beta_{0j}^{(2)} + \varepsilon_{ij}^{(2)} \\
\beta_{0j}^{(2)} &= \gamma_{00}^{(2)} + \gamma_{01}^{(2)} X_j + \mu_{0j}^{(2)} \\
Y_{ij} &= \beta_{0j}^{(3)} + \beta_{1j}^{(3)} M_{ij} + \varepsilon_{ij}^{(3)} \\
\beta_{0j}^{(3)} &= \gamma_{00}^{(3)} + \gamma_{01}^{(3)} X_j + \mu_{0j}^{(3)} \\
\beta_{1j}^{(3)} &= \gamma_{10}^{(3)}
\end{aligned}
\tag{9.2}
$$

就此模型而言，c 即$\gamma_{01}^{(1)}$；c' 即$\gamma_{01}^{(3)}$；a 即$\gamma_{01}^{(2)}$；b 即$\gamma_{10}^{(3)}$。

1-1-1模型：

$$
\begin{aligned}
Y_{ij} &= \beta_{0j}^{(1)} + \beta_{1j}^{(1)} X_{ij} + \varepsilon_{ij}^{(1)} \\
\beta_{0j}^{(1)} &= \gamma_{00}^{(1)} + \mu_{0j}^{(1)} \\
\beta_{1j}^{(1)} &= \gamma_{10}^{(1)} \\
M_{ij} &= \beta_{0j}^{(2)} + \beta_{1j}^{(2)} X_{ij} + \varepsilon_{ij}^{(2)} \\
\beta_{0j}^{(2)} &= \gamma_{00}^{(2)} + \mu_{0j}^{(2)} \\
\beta^{(2)}{}_{1j} &= \gamma_{10}^{(2)} \\
Y_{ij} &= \beta_{0j}^{(3)} + \beta_{1j}^{(3)} X_{ij} + \beta_{2j}^{(3)} M_{ij} + \varepsilon_{ij}^{(3)} \\
\beta_{0j}^{(3)} &= \gamma_{00}^{(3)} + \mu_{0j}^{(3)} \\
\beta_{1j}^{(3)} &= \gamma_{10}^{(3)} \\
\beta_{2j}^{(3)} &= \gamma_{20}^{(3)}
\end{aligned}
\tag{9.3}
$$

就此模型而言，c 即$\gamma_{10}^{(1)}$；c' 即$\gamma_{10}^{(3)}$；a 即$\gamma_{10}^{(2)}$；b 即$\gamma_{20}^{(3)}$。

9.2.2 多水平随机中介效应分析

多水平随机中介效应模型可通过在水平-2增加水平-1对应斜率方程的残差项获得（温忠麟等，2012：232-233）。

2-1-1模型：

$$
\begin{aligned}
Y_{ij} &= \beta_{0j}^{(1)} + \varepsilon_{ij}^{(1)} \\
\beta_{0j}^{(1)} &= \gamma_{00}^{(1)} + \gamma_{01}^{(1)} X_j + \mu_{0j}^{(1)} \\
M_{ij} &= \beta_{0j}^{(2)} + \varepsilon_{ij}^{(2)} \\
\beta_{0j}^{(2)} &= \gamma_{00}^{(2)} + \gamma_{01}^{(2)} X_j + \mu_{0j}^{(2)} \\
Y_{ij} &= \beta_{0j}^{(3)} + \beta_{1j}^{(3)} M_{ij} + \varepsilon_{ij}^{(3)} \\
\beta_{0j}^{(3)} &= \gamma_{00}^{(3)} + \gamma_{01}^{(3)} X_j + \mu_{0j}^{(3)} \\
\beta_{1j}^{(3)} &= \gamma_{10}^{(3)} + \mu_{1j}^{(3)}
\end{aligned}
\tag{9.4}
$$

1-1-1模型：

$$
\begin{aligned}
Y_{ij} &= \beta_{0j}^{(1)} + \beta_{1j}^{(1)} X_{ij} + \varepsilon_{ij}^{(1)} \\
\beta_{0j}^{(1)} &= \gamma_{00}^{(1)} + \mu_{0j}^{(1)} \\
\beta_{1j}^{(1)} &= \gamma_{10}^{(1)} + \mu_{1j}^{(1)} \\
M_{ij} &= \beta_{0j}^{(2)} + \beta_{1j}^{(2)} X_{ij} + \varepsilon_{ij}^{(2)} \\
\beta_{0j}^{(2)} &= \gamma_{00}^{(2)} + \mu_{0j}^{(2)} \\
\beta_{1j}^{(2)} &= \gamma_{10}^{(2)} + \mu_{1j}^{(2)} \\
Y_{ij} &= \beta_{0j}^{(3)} + \beta_{1j}^{(3)} X_{ij} + \beta_{2j}^{(3)} M_{ij} + \varepsilon_{ij}^{(3)} \\
\beta_{0j}^{(3)} &= \gamma_{00}^{(3)} + \mu_{0j}^{(3)} \\
\beta_{1j}^{(3)} &= \gamma_{10}^{(3)} + \mu_{1j}^{(3)} \\
\beta_{2j}^{(3)} &= \gamma_{20}^{(3)} + \mu_{2j}^{(3)}
\end{aligned}
\qquad (9.5)
$$

a、b、c、 c' 等系数与实际模型相应系数的对应关系与固定效应模型相同。

9.2.3　基于MLM的多水平中介效应分析存在的问题

基于MLM框架进行多水平中介效应分析中要注意变量中心化问题。如果不对多水平中介模型水平-1预测变量进行恰当的中心化，会引起组内效应（within-group effect）和组间效应（between-group effect）的混淆（Krull＆MacKinnon，2001；Zhang et al，2009；Preacher et al，2011）。

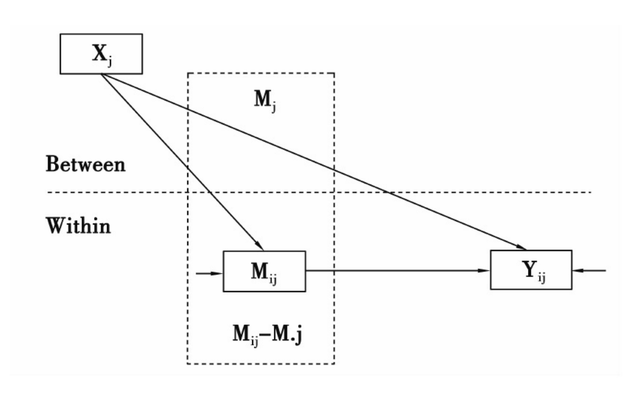


图9-4　传统MLM处理多水平中介效应（2-1-1模型）

比如，我们要研究学校学习氛围（X_j）通过学生阅读态度（M_{ij}）对阅读得分（Y_{ij}）的影响。对于这个2-1-1模型（见图9-4），实际上，在总的间接效应中，一部分是学校整体学习氛围对学校学生总体阅读态度（以均值呈现）有影响，进而影响学生阅读得分；另外一部分是学生个体在不同学校整体学习氛围中，阅读态度个体之间存在差异，导致个体阅读成绩不同。传统的MLM无法区分出这一点，造成混淆。

以固定斜率的2-1-1模型为例，式9.2属于混淆的（conflated）多水平回归模型。最后两个方程代入倒数第三个方程，得：

$$Y_{ij} = \gamma_{00}^{(3)} + \gamma_{01}^{(3)} X_j + \underbrace{\left[\gamma_{10}^{(3)} M_{ij}\right]}_{\text{混杂}} + \mu_{0j}^{(3)} + \varepsilon_{ij}^{(3)} \tag{9.6}$$

利用上述模型进行多水平分析，组间和组内效应混杂在一起。对中介变量做组均值中心化，可将中介效应的组内和组间部分区分开。

式9.2中，第三步对M_{ij}取组均值中心化，并在水平-2截距方程中增加组均值，得到不混淆的（unconflated）多水平回归模型：

$$Y_{ij} = \beta_{0j}^{(3)} + \beta_{1j}^{(3)}(M_{ij} - M_{.j}) + \varepsilon_{ij}^{(3)}$$
$$\beta_{0j}^{(3)} = \gamma_{00}^{(3)} + \gamma_{01}^{(3)} X_j + \gamma_{02}^{(3)} M_{.j} + \mu_{0j}^{(3)} \quad (9.7)$$
$$\beta_{1j}^{(3)} = \gamma_{10}^{(3)}$$

最后两个方程代入倒数第三个方程，

$$Y_{ij} = \gamma_{00}^{(3)} + \gamma_{01}^{(3)} X_j + \underbrace{[\gamma_{02}^{(3)}]}_{\text{组间}} M_{.j} + \underbrace{[\gamma_{10}^{(3)}]}_{\text{组内}} (M_{ij} - M_{.j}) + \mu_{0j}^{(3)} + \varepsilon_{ij}^{(3)} \quad (9.8)$$

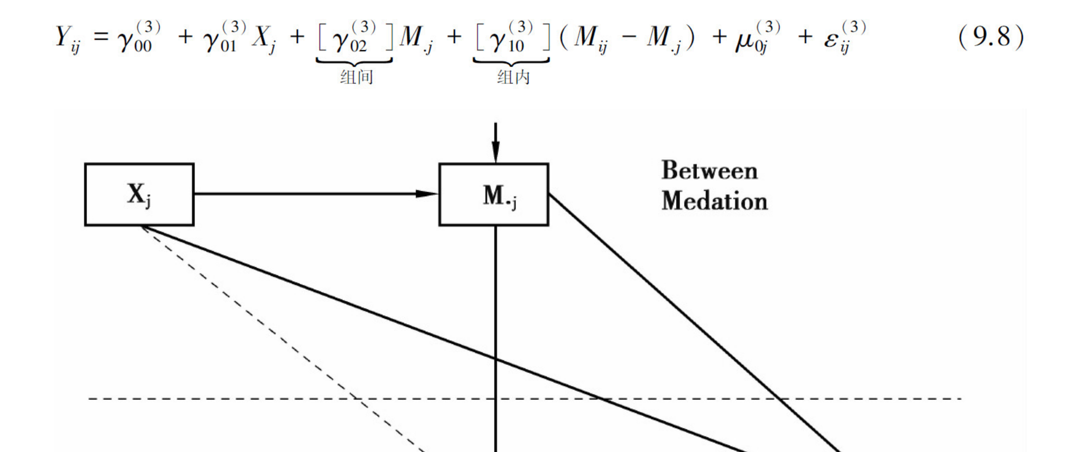


图9-5　组中心化的MLM处理多水平中介效应（2-1-1）

这样，经过组中心化，区分出了组间中介效应。实际上，在2-1-1模型中，因为X_j只有组间层面的方差，只能在水平-2组间变化。对某个固定的学校j而言，X_j是个定值，其对水平-1M_{ij}并不能产生直接的影响，而是通过影响第ij个个体所在群组的均值即$M_{.j}$间接实现的。因此，路径a实际上属于水平-2。所以如果X_j只是影响M_{ij}的组间部分，在1-1关系中M_{ij}对Y_{ij}组内效应不应该包含在2-1-1的中介因果链条中。

在2-1-1模型中，组内个体围绕组均值的变化应视为随机，组内中介效应意义不大。基于组中心化的2-1-1模型，多水平组间中介效应可以以$\gamma_{01}^{(1)}-\gamma_{02}^{(3)}$ 或$\gamma_{01}^{(2)}\times\gamma_{02}^{(3)}$ 进行检验。

对于2-2-1模型，组均值中心化和总均值中心化（或直接使用原始变量形式）等价，中介效应不存在混淆的问题。对于式9.1所刻画的2-2-1模型，多水平组间中介可由$\gamma_{01}^{(1)}$ - $\gamma_{01}^{(3)}$ 或$\gamma_{01}^{(2)}\times\gamma_{02}^{(3)}$ 进行检验。

在2-1-1模型，组均值中心化和总均值中心化不等价，应使用组均值中心化计算组水平中介效应。对M_{ij} 总均值中心化，暗含了较强的假设。对于式9.2，最后两个方程代入倒数第三个方程，整理得：

$$\begin{aligned} Y_{ij} &= \gamma_{00}^{(3)} + \gamma_{01}^{(3)}X_j + \gamma_{10}^{(3)}M_{ij} + \mu_{0j}^{(3)} + \varepsilon_{ij}^{(3)} \\ &= \gamma_{00}^{(3)} + \gamma_{01}^{(3)}X_j + \underbrace{[\gamma_{10}^{(3)}]}_{\text{组间}}M_{..} + \underbrace{[\gamma_{10}^{(3)}]}_{\text{组内}}(M_{ij} - M_{..}) + \mu_{0j}^{(3)} + \varepsilon_{ij}^{(3)} \end{aligned} \tag{9.9}$$

也就是说，假定了组间和组内效应相同，与式9.8比较，即强制$\gamma_{02}^{(3)}=\gamma_{10}^{(3)}$ 。

模拟研究（Zhang et al，2009）表明，在2-1-1模型中，使用总均值中心化犯第Ⅰ种类型错误的概率增加，在组内部分（1-1）存在或强而组间部分不存在或弱的时候，对不同层面的中介效应不加区分，整体高估中介效应，结果也会导致对组间效应的高估。而此时研究者可能更关心组间效应而不关注组内效应。为了避免此类错误，多水平中介效应分析（2-1-1、1-1-1模型）中，检验中介效应必须区分组间和组内不同的成分。

在2-1-1模型中，一般都是关注组间中介效应。而在1-1-1模型中，情况更为复杂，可就是否存在组间的中介效应、组内中介效应，

或就两种中介效应同时进行研究甚至进行大小的比较。如果将组间和组内部分混为一谈，会导致水平-1、水平-2或两个层面中介效应的误识（Zhang et al，2009）。

对于1-1-1模型，组均值中心化和总均值中心化不等价，应做组均值中心化来区分组水平中介效应和个体水平中介效应。按传统的方法，是就式9.3中$\gamma_{10}^{(1)}-\gamma_{10}^{(3)}$ 或$\gamma_{10}^{(2)}\times\gamma_{20}^{(3)}$ 进行检验。但这样混淆了组内和组间中介效应，会导致水平-1、水平-2或两个水平同时中介效应的错误估计。

为了避免这个问题，可在1-1-1模型中，同时对X_{ij} 和M_{ij} 进行组中心化，并在水平-2（针对水平-1截距即β_{0j} 的）方程中增加相应的组均值：

$$
\begin{aligned}
Y_{ij} &= \beta_{0j}^{(1)} + \beta_{1j}^{(1)}(X_{ij} - X_{.j}) + \varepsilon_{ij}^{(1)} \\
\beta_{0j}^{(1)} &= \gamma_{00}^{(1)} + \gamma_{01}^{(1)} X_{.j} + \mu_{0j}^{(1)} \\
\beta_{1j}^{(1)} &= \gamma_{10}^{(1)} \\
M_{ij} &= \beta_{0j}^{(2)} + \beta_{1j}^{(2)}(X_{ij} - X_{.j}) + \varepsilon_{ij}^{(2)} \\
\beta_{0j}^{(2)} &= \gamma_{00}^{(2)} + \gamma_{01}^{(2)} X_{.j} + \mu_{0j}^{(2)} \\
\beta_{1j}^{(2)} &= \gamma_{10}^{(2)} \\
Y_{ij} &= \beta_{0j}^{(3)} + \beta_{1j}^{(3)}(X_{ij} - X_{.j}) + \beta_{2j}^{(3)}(M_{ij} - M_{.j}) + \varepsilon_{ij}^{(3)} \\
\beta_{0j}^{(3)} &= \gamma_{00}^{(3)} + \gamma_{01}^{(3)} X_{.j} + \gamma_{02}^{(3)} M_{.j} + \mu_{0j}^{(3)} \\
\beta_{1j}^{(3)} &= \gamma_{10}^{(3)} \\
\beta_{2j}^{(3)} &= \gamma_{20}^{(3)}
\end{aligned}
\qquad (9.10)
$$

通过整理，得：

$$
\begin{aligned}
Y_{ij} &= \gamma_{00}^{(1)} + \gamma_{01}^{(1)} X_{.j} + \gamma_{10}^{(1)} (X_{ij} - X_{.j}) + \mu_{0j}^{(1)} + \varepsilon_{ij}^{(1)} \\
M_{ij} &= \gamma_{00}^{(2)} + \gamma_{01}^{(2)} X_{.j} + \gamma_{10}^{(2)} (X_{ij} - X_{.j}) + \mu_{0j}^{(2)} + \varepsilon_{ij}^{(2)} \\
Y_{ij} &= \gamma_{00}^{(3)} + \gamma_{01}^{(3)} X_{.j} + \gamma_{02}^{(3)} M_{.j} + \gamma_{10}^{(3)} (X_{ij} - X_{.j}) + \\
&\quad \gamma_{20}^{(3)} (M_{ij} - M_{.j}) + \mu_{0j}^{(3)} + \varepsilon_{ij}^{(3)}
\end{aligned}
\tag{9.11}
$$

就上式组中心化处理的1-1-1模型，组内中介效应用$\gamma_{10}^{(1)}-\gamma_{10}^{(3)}$ 或$\gamma_{10}^{(2)}\times\gamma_{20}^{(3)}$ 来检验；组间中介效应可用$\gamma_{01}^{(1)}-\gamma_{01}^{(3)}$ 或$\gamma_{01}^{(2)}\times\gamma_{02}^{(3)}$ 来检验。

基于Lüdtke et al（2008）、Preacher et al（2010；2011 ）、Preacher（2011 ）等研究，Preacher（2015）总结了基于MLM的中介效应分析的局限：（1）2-1斜率其实对水平-2而言是特定的，非跨层的斜率；（2）1-1效应是混淆的，除非对水平-1测量变量进行中心化，但即使通过组均值中心化将混淆的效应分解为组特定的成分，以观测的组均值作为预测变量，水平-2效应仍然存在偏误；（3）MLM不能处理结果变量处于高层的模型。相比之下，基于MSEM的多水平中介效应分析建模方式，更为富于弹性，估计结果也更具信度。

9.3　基于MSEM的多水平中介效应分析

9.3.1　基于MSEM的多水平中介效应分析的优势

尽管通过组中心化，MLM方法区分出组内和组间中介效应，但仍然存在局限。MLM方法实际上是将观测的某个预测变量的组均值

（observed means）作为该变量在组群层面潜在特征的代理，这样做信度较差（题项过少会加剧这种情况），会造成估计的偏差，包括：1-1关系组群层面主效应偏差（Lüdtke et al，2008），以及2-1-1模型中组间中介效应偏差（Preacher et al，2010）。

从MSEM来看，很容易理解这个问题。在SEM的框架下，测量模型部分多个指标和一个潜变量对应的结构，可以允许测量误差校正。在MSEM中，除了多测量误差的校正，水平-1单位可以被理解为水平-2潜在构念的指标（所谓“人也是变量”，参见Mehta＆Neale，2005），从而实现对抽样误差的校正。正像因子分析当中那样，指标过少，或缺少共同度，可以造成因子之间关系估计的偏差，水平-1单位少或者ICC过小，也就是缺乏信度的情况下，同样会造成组间效应估计的偏误（Lüdtke et al，2008；Neuhaus＆McCulloch，2006）。MSEM的优势在于，可将水平-1变量含有的无法直接观测的组群成分作为潜变量（latent means），从而准确估计出抽样误差大小，提高测量的精度。正是因为多水平结构方程模型实现了对测量误差和抽样误差的校正，有学者（Marsh et al，2012）将之称作“双重潜多水平模型”。

模拟研究（Preacher et al，2011；Li＆Beretvas，2013）结果表明，在处理多水平中介效应问题上，MSEM相对于MLM而言，不仅能有效减少中介效应估计的偏差，而且在中介效应的95％置信区间覆盖率指标上表现也更优。MSEM进行多水平中介分析时，可以提供模型拟合指数，从而可以对模型的拟合程度进行评价。

另外，由于MLM的定义方式限制，只能分析因变量在水平-1的多水平中介效应，不能处理2-1-2、1-2-1等特殊的多水平中介效应模型设计。而MSEM的方式对因变量在水平-1和水平-2的多水平中介效应均可以分析，因而具有更广的适用性。

总的来看，在MSEM中，变量分为3种类型：①只在水平-2变化的变量，也就是组群层面变量；②只在水平-1变化的变量，也就是个体层面的变量，ICC微不足道（一般小于0.05）；③在两个层面都变化的变量（一般都是基于水平-1进行测量）。正如第8章第1节 所述，MSEM使用的是一种“分解优先”的思路。按照MLC的策略，第1、2类变量，各自在所属水平直接参与构建方程；对于第3类变量，自动分解为组内部分和组间两部分，分别在两个水平参与建模。这样，在多水平中介效应分析时，通过MSEM方法，对于一些属于第3类型的水平-1变量，*Mplus* 会自动生成该变量的组间部分，参与水平-2的中介效应分析。这种处理方式相当于将所有水平-1变量均视作带有随机截距，而不仅是结果变量Y。

模型以图形来表示，见图9-6、图9-7、图9-8所示，与基于MLM的2-2-1、2-1-1和1-1-1多水平中介效应模型图9-1、图9-2、图9-3进行比照可以看到，基于MSEM的多水平中介效应分析中，不同水平只考虑本层要素的方差，其间也没有代表因果作用路径的跨层箭头存在。各水平之间是互相独立的。

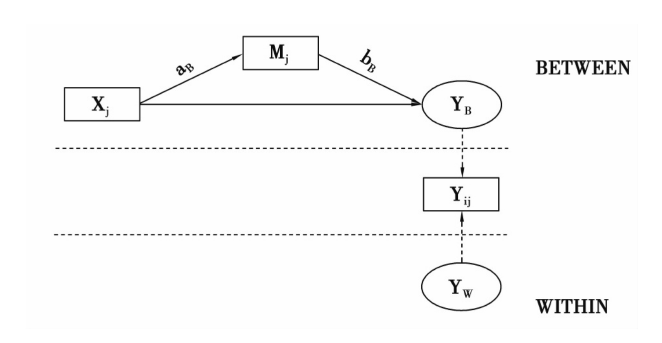


图9-6　基于MSEM的2-2-1模型中介效应

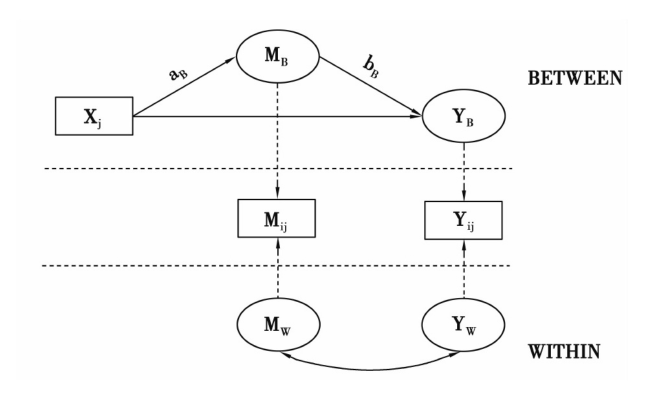


图9-7　基于MSEM的2-1-1模型中介效应

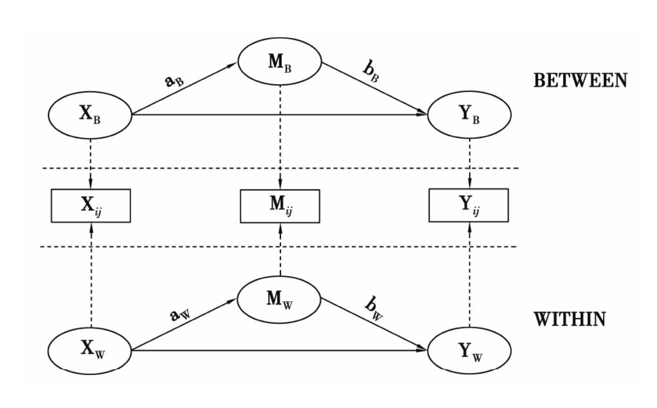


图9-8　基于MSEM的1-1-1模型中介效应

基于“分解优先”的策略，MSEM多水平中介效应分析框架，层次清晰，不会混淆组内和组间效应。水平-2模型解释水平-2要素的方差，包括水平-2变量及水平-1变量的组间部分；水平-1模型解释水平-1要素的方差，包括水平-1变量的组内部分，以及严格界定只在水平-1变化无组间方差的变量。

由此而言，说MSEM框架中的“多水平中介效应分析”似乎有些“误导”：在基于MSEM的多水平中介效应分析的模型中，可能有组间中介效应，也可能有组内中介效应，但并不存在从水平-2到水平-1的“跨层”中介效应。比如，从图9-7可以看到，区分组内和组间部分后，2-1-1模型其实也是一个组间中介效应模型；从图9-8来看，1-1-1模型组内和组间均有中介效应。实际上，基于MSEM的多水平中介效应

分析的一个原则是，只要预测变量、中介变量和结果变量中有一个属于水平-2变量，则中介效应必定在组间水平发生（Preacher et al，2010）。

9.3.2 以矩阵表达的基于MSEM多水平中介效应分析

对于上述基于MSEM的多水平中介效应模型，以矩阵形式表达模型设定更为直观（参见Preacher et al，2010）。

（1）2-2-1中介效应模型

测量方程：

$$
\begin{aligned}
Y_{ij} &= \Lambda \eta_{ij} \\
&= \begin{bmatrix} X_j \\ M_j \\ Y_{ij} \end{bmatrix} = \left[\begin{array}{c:ccc} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} \eta_{Yij} \\ \hdashline \eta_{Xj} \\ \eta_{Mj} \\ \eta_{Yj} \end{bmatrix}
\end{aligned}
\quad (9.12)
$$

组内结构方程：

$$\eta_{ij} = \alpha_j + \zeta_{ij}$$

$$= \begin{bmatrix} \eta_{Yij} \\ ---- \\ \eta_{Xj} \\ \eta_{Mj} \\ \eta_{Yj} \end{bmatrix} = \begin{bmatrix} 0 \\ ---- \\ \alpha_{\eta Xj} \\ \alpha_{\eta Mj} \\ \alpha_{\eta Yj} \end{bmatrix} + \begin{bmatrix} \zeta_{Yij} \\ ---- \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{9.13}$$

组间结构方程：

$$\eta_j = \mu + \beta\eta_j + \zeta_j$$

$$= \begin{bmatrix} \alpha_{\eta Xj} \\ \alpha_{\eta Mj} \\ \alpha_{\eta Yj} \end{bmatrix} = \begin{bmatrix} \mu_{\alpha\eta Xj} \\ \mu_{\alpha\eta Mj} \\ \mu_{\alpha\eta Yj} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ \beta_{MX} & 0 & 0 \\ \beta_{YX} & \beta_{YM} & 0 \end{bmatrix} \begin{bmatrix} \alpha_{\eta Xj} \\ \alpha_{\eta Mj} \\ \alpha_{\eta Yj} \end{bmatrix} + \begin{bmatrix} \zeta_{\alpha\eta Xj} \\ \zeta_{\alpha\eta Mj} \\ \zeta_{\alpha\eta Yj} \end{bmatrix} \tag{9.14}$$

对于2-2-1模型，只有1个变量在组内层面，因此无组内间接效应，β矩阵由组间间接效应构成：

$$\beta = \begin{bmatrix} 0 & 0 & 0 \\ \beta_{MX} & 0 & 0 \\ \beta_{YX} & \beta_{YM} & 0 \end{bmatrix} \tag{9.15}$$

X_j 经由M_j 对Y_{ij} 总间接效应为：

$$F = (I - \beta)^{-1} - I - \beta = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \beta_{MX}\beta_{YM} & 0 & 0 \end{bmatrix} \tag{9.16}$$

其中，β_{MX} β_{YM} 也就是组间间接效应。

（2）2-1-1中介效应模型，固定斜率

测量方程：

$$Y_{ij} = \Lambda \eta_{ij}$$

$$= \begin{bmatrix} X_j \\ M_{ij} \\ Y_{ij} \end{bmatrix} = \left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} \eta_{Mij} \\ \eta_{Yij} \\ \hdashline \eta_{Xj} \\ \eta_{Mj} \\ \eta_{Yj} \end{bmatrix} \tag{9.17}$$

组内结构方程：

$$\eta_{ij} = \alpha_j + B_j \eta_{ij} + \zeta_{ij}$$

$$= \begin{bmatrix} \eta_{Mij} \\ \eta_{Yij} \\ \hdashline \eta_{Xj} \\ \eta_{Mj} \\ \eta_{Yj} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \hdashline \alpha_{\eta Xj} \\ \alpha_{\eta Mj} \\ \alpha_{\eta Yj} \end{bmatrix} + \left[\begin{array}{cc|ccc} 0 & 0 & 0 & 0 & 0 \\ B_{YMj} & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] + \begin{bmatrix} \zeta_{Mij} \\ \zeta_{Yij} \\ \hdashline 0 \\ 0 \\ 0 \end{bmatrix} \tag{9.18}$$

组间结构方程：

$$\eta_j = \mu + \beta \eta_j + \zeta_j$$

$$= \begin{bmatrix} \alpha_{\eta Xj} \\ \alpha_{\eta Mj} \\ \alpha_{\eta Yj} \end{bmatrix} = \begin{bmatrix} \mu_{\alpha\eta Xj} \\ \mu_{\alpha\eta Mj} \\ \mu_{\alpha\eta Yj} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ \beta_{MX} & 0 & 0 \\ \beta_{YX} & \beta_{YM} & 0 \end{bmatrix} \begin{bmatrix} \alpha_{\eta Xj} \\ \alpha_{\eta Mj} \\ \alpha_{\eta Yj} \end{bmatrix} + \begin{bmatrix} \zeta_{BYMj} \\ \zeta_{\alpha\eta Xj} \\ \zeta_{\alpha\eta Mj} \end{bmatrix} \tag{9.19}$$

（3）2-1-1中介效应模型，随机斜率

测量方程：

$$Y_{ij} = \Lambda\eta_{ij}$$

$$= \begin{bmatrix} X_j \\ M_{ij} \\ Y_{ij} \end{bmatrix} = \left[\begin{array}{cc:ccc} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} \eta_{Mij} \\ \eta_{Yij} \\ \hdashline \eta_{Xj} \\ \eta_{Mj} \\ \eta_{Yj} \end{bmatrix} \tag{9.20}$$

组内结构方程：

$$\eta_{ij} = \alpha_j + \mathrm{B}_j\eta_{ij} + \zeta_{ij}$$

$$= \begin{bmatrix} \eta_{Mij} \\ \eta_{Yij} \\ \hdashline \eta_{Xj} \\ \eta_{Mj} \\ \eta_{Yj} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \hdashline \alpha_{\eta Xj} \\ \alpha_{\eta Mj} \\ \alpha_{\eta Yj} \end{bmatrix} + \left[\begin{array}{cc:ccc} 0 & 0 & 0 & 0 & 0 \\ B_{YMj} & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] + \begin{bmatrix} \zeta_{Mij} \\ \zeta_{Yij} \\ \hdashline 0 \\ 0 \\ 0 \end{bmatrix} \tag{9.21}$$

组间结构方程：

$$\eta_j = \mu + \beta\eta_j + \zeta_j$$

$$= \begin{bmatrix} B_{YMj} \\ \hdashline \alpha_{\eta Xj} \\ \alpha_{\eta Mj} \\ \alpha_{\eta Yj} \end{bmatrix} = \begin{bmatrix} \mu_{BYMj} \\ \hdashline \mu_{\alpha\eta Xj} \\ \mu_{\alpha\eta Mj} \\ \mu_{\alpha\eta Yj} \end{bmatrix} + \left[\begin{array}{c:ccc} 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 \\ 0 & \beta_{MX} & 0 & 0 \\ 0 & \beta_{YX} & \beta_{YM} & 0 \end{array}\right] \begin{bmatrix} B_{YMj} \\ \hdashline \alpha_{\eta Xj} \\ \alpha_{\eta Mj} \\ \alpha_{\eta Yj} \end{bmatrix} + \begin{bmatrix} \zeta_{BYMj} \\ \hdashline \zeta_{\alpha\eta Xj} \\ \zeta_{\alpha\eta Mj} \\ \zeta_{\alpha\eta Yj} \end{bmatrix} \tag{9.22}$$

在2-1-1模型中，有两个组内层面变量，但仍然没有组内间接效应。X_j 经由M_{ij} 对Y_{ij} 总（组间）间接效应包含在 β 的3×3的子矩阵之中。总的间接效应：

$$F = (I-\beta_B)^{-1}-I-\beta_B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \beta_{MX}\beta_{YM} & 0 & 0 \end{bmatrix} \qquad (9.23)$$

其中，β_{MX} β_{YM} 即组间间接效应。

（4）1-1-1中介效应模型，随机斜率

测量方程：

$$Y_{ij} = \Lambda\eta_{ij}$$

$$= \begin{bmatrix} X_{ij} \\ M_{ij} \\ Y_{ij} \end{bmatrix} = \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} \eta_{Xij} \\ \eta_{Mij} \\ \eta_{Yij} \\ \hdashline \eta_{Xj} \\ \eta_{Mj} \\ \eta_{Yj} \end{bmatrix} \qquad (9.24)$$

组内结构方程：

$$\eta_{ij} = \alpha_j + \mathrm{B}_j\eta_{ij} + \zeta_{ij}$$

$$= \begin{bmatrix} \eta_{Xij} \\ \eta_{Mij} \\ \eta_{Yij} \\ \hdashline \eta_{Xj} \\ \eta_{Mj} \\ \eta_{Yj} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \hdashline \alpha_{\eta Xj} \\ \alpha_{\eta Mj} \\ \alpha_{\eta Yj} \end{bmatrix} + \left[\begin{array}{ccc:ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ B_{MXj} & 0 & 0 & 0 & 0 & 0 \\ B_{YXj} & B_{YMj} & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{bmatrix} \eta_{Xij} \\ \eta_{Mij} \\ \eta_{Yij} \\ \hdashline \eta_{Xj} \\ \eta_{Mj} \\ \eta_{Yj} \end{bmatrix} + \begin{bmatrix} \zeta_{Xij} \\ \zeta_{Mij} \\ \zeta_{Yij} \\ \hdashline 0 \\ 0 \\ 0 \end{bmatrix} \qquad (9.25)$$

组间结构方程：

$$\eta_j = \mu + \beta\eta_j + \zeta_j$$

$$= \begin{bmatrix} B_{MXj} \\ B_{YXj} \\ B_{YMj} \\ \hdashline \alpha_{\eta Xj} \\ \alpha_{\eta Mj} \\ \alpha_{\eta Yj} \end{bmatrix} = \begin{bmatrix} \mu_{BMXj} \\ \mu_{BYXj} \\ \mu_{BYMj} \\ \hdashline \mu_{\alpha\eta Xj} \\ \mu_{\alpha\eta Mj} \\ \mu_{\alpha\eta Yj} \end{bmatrix} + \left[\begin{array}{ccc:ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \beta_{MX} & 0 & 0 \\ 0 & 0 & 0 & \beta_{YX} & \beta_{YM} & 0 \end{array}\right] \begin{bmatrix} B_{MXj} \\ B_{YXj} \\ B_{YMj} \\ \hdashline \alpha_{\eta Xj} \\ \alpha_{\eta Mj} \\ \alpha_{\eta Yj} \end{bmatrix} + \begin{bmatrix} \zeta_{BMXj} \\ \zeta_{BYXj} \\ \zeta_{BYMj} \\ \hdashline \zeta_{\alpha\eta Xj} \\ \zeta_{\alpha\eta Mj} \\ \zeta_{\alpha\eta Yj} \end{bmatrix} \qquad (9.26)$$

在1-1-1模型中，X_{ij} 、M_{ij} 和Y_{ij} 都有组内和组间的变化，组内和组间的间接效应都可估计（Preacher et al，2010）。X_{ij} 经由M_{ij} 对Y_{ij} 组间间接效应包含在β的3 ×3的子矩阵之中。组间间接效应：

$$\mathrm{F} = (\mathrm{I}\text{-}\beta_{\mathrm{B}})^{-1}\text{-}\mathrm{I}\text{-}\beta_{\mathrm{B}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \beta_{MX}\beta_{YM} & 0 & 0 \end{bmatrix} \qquad (9.27)$$

其中，β_{MX} β_{YM} 即组间间接效应，即X_{ij} 组间部分经由M_{ij} 组间部分对Y_{ij} 组间部分的间接效应。

另外，β 也包含构成组内间接效应的系数，这些元素既可以是固定斜率，也可以是随机斜率的均值。设定B_{MXj} 和B_{YMj} 固定或随机将影响组内间接效应的值。如果二者都是随机斜率均值，组内间接效应则是：$\mu_{BMXj}\mu_{BYMj}+\Psi_{\zeta BMXj,\ \zeta BYXj}$。其中，$\Psi_{\zeta BMXj,\ \zeta BYXj}$ 为组群水平随机斜率B_{MXj} 和B_{YMj} 的协方差。如果B_{MXj} 和B_{YMj} 固定，组内简介效应就是：$\mu_{BMXj}\mu_{BYMj}$。

（5）小结

在MSEM框架下，2-2-1模型（组间）中介效应可以通过$a_B\times b_B$ 来检验。对于2-1-1模型固定效应，（组间）中介效应通过$a_B\times b_B$ 检验。对于带有随机斜率的2-1-1模型，b 分为组间和组内两部分，一部分是水平-2 M_j 对Y_j 的系数b_B ，另一部分是水平-1 M_{ij} 对Y_{ij} 的系数b_W 。由于随机变化，b_W 取均值作为斜率的估计（即随机斜率的均值，实际表现于水平-2），总的间接效应即$a_B\times$（b_B+b_W）。对于M_{ij} 、Y_{ij} 对应方程均固定斜率的1-1-1模型，组内中介效应通过$a_W\times b_W$ 检验，组间中介效应通过$a_B\times b_B$ 检验。对于斜率均随机变化的1-1-1模型，组内间接效应为$a_W\times b_W$ +Cov（a_W，b_W）（Kenny et al，2003；Preacher et al，2010）。

表9-2　基于MSEM的两水平中介效应表达式

模　型	固定/随机斜率	组间中介效应	组内中介效应
2-2-1	无	$a_B \times b_B$	无
2-1-1	$M_{ij} \rightarrow Y_{ij}$固定斜率	$a_B \times b_B$	无
2-1-1	$M_{ij} \rightarrow Y_{ij}$随机斜率	$a_B \times (b_B + b_W)$	无
1-1-1	$X_{ij} \rightarrow M_{ij}$、$M_{ij} \rightarrow Y_{ij}$固定斜率	$a_B \times b_B$	$a_W \times b_W$
1-1-1	$X_{ij} \rightarrow M_{ij}$、$M_{ij} \rightarrow Y_{ij}$随机斜率	$a_B \times b_B$	$a_W \times b_W + Cov(a_W, b_W)$

中介效应分析中，由于乘积项$a \times b$通常不服从正态分布，推荐使用自举法（Monte Carlo bootstrapping）进行中介效应检验（Lockwood & MacKinnon，1998；MacKinnon，2007；Preacher & Selig，2012）。不过目前在M*plus*中，多水平模型并不支持使用自举法，也就是在ANALYSIS部分使用TYPE＝TWOLEVEL的同时进行bootstrap设定。OUTPUT部分设定CINTERVAL属于输出基于正态假设的置信区间。如果需要输出自举法汇报的经验分布，可以根据M*plus*输出结果中a、b路径系数及相应标准误等参数，另行编写程序进行以参数为基础的自举法（parameter-based bootstrapping）估计复合系数的置信区间（参见刘东等，2012）[3]。

9.4　多水平中介效应分析实例

参照Preacher et al（2010）提出的多水平中介效应分析的基本规则和范例，在不考虑理论意义和相关变量（显／潜）聚合意义的情况下，作为示例，以下使用实际数据进行多水平中介效应分析。

9.4.1　常见模型

多水平中介效应分析时，2-2-1、2-1-1、1-1-1模型在实际中最为常见。由于涉及模型较多，为节省篇幅起见，这里只提供了相关代码，未显示OUTPUT原始输出结果，主要结果汇总于后续表格中。

（1）2-2-1模型

PISA数据中，学生行为studbeha（M）、教师行为teacbeha（X）属于水平-2变量，以下研究其与水平-1学生数学成绩pvmath1（Y）的关系。相关代码见框9-1到框9-2。

框9-1　2-2-1模型单水平中介效应分析

```
TITLE: Simple Mediation Analysis;
DATA: File is 221. dat;
Variable:
  Names are
    schoolid pv1math studbeha teacbeha;
  Missing are all（－9999）;
  USEVARIABLES＝pv1math studbeha teacbeha;
ANALYSIS:
  BOOTSTRAP IS  5000;
MODEL:
  studbeha ON teacbeha（a）;   ！m对x回归，斜率标记为a
  pv1math ON studbeha（b）;   ！y对m回归，斜率标记为b
  pv1math ON teacbeha;
```

```
MODEL CONSTRAINT:
  NEW（ab）;                        ! 命名间接效应
  ab＝a*b;                          ! 计算间接效应
OUTPUT: CINTERVAL（BCBOOTSTRAP）;
```

框9-2　基于MSEM的2-2-1模型多水平中介效应分析

```
TITLE: 2-2-1 Mediation（MSEM）
DATA: File is 221. dat;
Variable:
  Names are
    schoolid pv1math studbeha teacbeha;
  Missing are all（－9999）;
  USEVARIABLES＝pv1math studbeha teacbeha;
  BETWEEN＝studbeha teacbeha;
  CLUSTER＝schoolid;
ANALYSIS:
  TYPE＝TWOLEVEL;
MODEL:
  %WITHIN%
  pv1math;
  %BETWEEN%
  studbeha ON teacbeha（a）;          ! m对x回归，斜率标记为a
  pv1math ON studbeha（b）;          ! y对m回归，斜率标记为b
  pv1math ON teacbeha;
```

```
MODEL CONSTRAINT:
  NEW（ab）;                              ! 命名（组间）间接效应
  ab＝a*b;                                ! 计算（组间）间接效应
OUTPUT: CINTERVAL;
```

表9-3　中介效应2-2-1模型结果

模　型	位　置	a	b	间接效应[95%置信区间]
单水平	未区分	0.630	20.190	12.725[10.167，15.125]
MSEM	组间	0.602	20.036	12.060[5.246，18.873]

在2-2-1模型中，中介效应其实是组间间接效应。从结果来看，单水平与多水平模型输出结果中，间接效应的点估计比较接近，但95%置信区间相差比较大。可以看到，由于没有考虑嵌套结构，使用单水平模型相对于MSEM会低估标准误。

框9-3给出的是使用R软件利用框9-2模型相关输出结果对中介效应经验分布的自举法估计。该方法估计的中介效应95%置信区间为［5.031，19.929］，比使用正态分布假定下的结果略大。对于一般的多水平中介效应分析，均可考虑采取这一方式获得自举法估计的中介效应经验分布。

框9-3　基于MSEM的2-2-1模型中介效应经验分布R代码

```
#a为水平-2　X对M的预测效应                      #
#astd为a的标准误                                #
#b为水平-2　M对Y的预测效应                      #
```

```
#bstd为b的标准误                                         #
# " rep＝20000 " 定义了重取样的次数为20000   #
# " conf＝95 " 定义了95％的CI                     #

a＝0.602
b＝20.036
astd＝0.077
bstd＝5.795
rep＝20000
conf＝95
avec＝rnorm（rep）*astd＋a
bvec＝rnorm（rep）*bstd＋b
ab＝avec*bvec
low＝（1－conf／100）／2
upp＝（（1－conf／100）／2）＋（conf／100）
LL＝quantile（ab，low）
UL＝quantile（ab，upp）
LL4＝format（LL，digits＝5）
UL4＝format（UL，digits＝5）
hist（ab，breaks＝'FD'，col＝'skyblue'，xlab＝paste（conf，
% Confidence  Interval'，'LL'，LL4，'UL'，UL4），main
＝'Distribution of Indirect Effect'）
```

（2）2-1-1模型

以水平-2学校教育资源质量scmatedu（X），水平-1阅读态度readatt（M）和阅读成绩readsco（Y）的关系为例，构建相关模型。相关代码见框9-4至框9-8。

框9-4　2-1-1模型单水平中介效应分析

```
TITLE：Simple Mediation Analysis;
DATA：File is 211. dat;
Variable:
  Names are
    schoolid scmatedu readatt readsco greadatt;
  Missing are all（－9999）;
  USEVARIABLES＝readsco readatt scmatedu;
ANALYSIS:
  BOOTSTRAP IS 5000;
MODEL:
  readAtt ON scmatedu（a）;        ！m对x回归，标记为a
  readSco ON readAtt（b）;          ！y对m回归，标记为b
  readSco ON scmatedu;
MODEL CONSTRAINT:
  NEW（indb）;                     ！命名间接效应
  indb＝a*b;                       ！计算间接效应
OUTPUT：CINTERVAL（BCBOOTSTRAP）;
```

框9-5　2-1-1基于传统MLM的多水平中介效应分析

```
TITLE: 2-1-1 mediation (traditional MLM)
DATA: File is 211. dat;
Variable:
  Names are
    schoolid scmatedu readatt readsco greadatt;
  Missing are all (－9999);
  USEVARIABLES＝readsco readatt scmatedu;
  BETWEEN＝scmatedu;
  CLUSTER＝schoolid;
ANALYSIS:
  TYPE＝TWOLEVEL;
MODEL:
  %WITHIN%
   readsco readatt;                     ！估计y和m水平-1（残差）协方差
   readsco ON readatt (b);          ！y对m回归，斜率记作b
  %BETWEEN%
   scmatedu readatt readsco;          ！估计x、m、y水平-2（残差）协方差
   readatt ON scmatedu (a);        ！m对x回归，斜率记作a
   readsco ON readatt (b);        ！y对m回归，约束斜率等于b
   readsco ON scmatedu;
MODEL CONSTRAINT:
  NEW (indb);                 ！命名间接效应
  indb＝a*b;                 ！计算间接效应
OUTPUT: CINTERVAL;
```

框9-6　基于Unconflated MLM的2-1-1模型多水平中介效应分析

```
TITLE：2-1-1 mediation（unconflated MLM）
DATA：File is 211. dat；
Variable：
  Names are
    schoolid scmatedu readatt readsco greadatt；
  Missing are all（－9999）；
  USEVARIABLES＝scmatedu readatt greadatt readsco；
  BETWEEN＝scmatedu greadatt；
  WITHIN＝readatt；
  CLUSTER＝schoolid；
DEFINE：CENTER readatt（GROUPMEAN）
ANALYSIS：
  TYPE＝TWOLEVEL；
MODEL：
  %WITHIN%
   readsco readatt；                    ！估计y和m水平-1（残差）协方差
   readsco ON readatt；
   [readatt@0]                          ！readatt为组中心化值，所以固定均值为0
  %BETWEEN%
   greadatt readsco；                   ！估计x、m、y水平-2（残差）协方差
```

```
   greadatt ON scmatedu（a）；      ！m对x回归，标记为a
   readsco ON greadatt（b）；        ！y对m回归，标记为b
   readsco ON scmatedu;
MODEL CONSTRAINT:
  NEW（indb）；                    ！命名（组间）间接效应
  indb＝a*b;                     ！计算（组间）间接效应
OUTPUT: CINTERVAL;
```

框9-7　基于MSEM的2-1-1模型多水平中介效应分析

```
TITLE: 2-1-1 mediation（MSEM）
DATA: File is 211. dat;
Variable:
  Names are
    schoolid scmatedu readatt readsco greadatt;
  Missing are all（－9999）；
  USEVARIABLES＝scmatedu readatt readsco;
  BETWEEN＝scmatedu;
  CLUSTER＝schoolid;
ANALYSIS:
  TYPE＝TWOLEVEL;
MODEL:
  %WITHIN%
   readsco readatt;             ！估计y和m水平-1（残差）协方差
   readsco ON readatt;
```

```
  %BETWEEN%
   scmatedu readatt readsco;      ！估计x、m、y水平-2（残差）协方差
   readatt ON scmatedu（a）；      ！m对x回归，标记为a
   readsco ON readatt（b）；      ！y对m回归，标记为b
   readsco ON scmatedu;
MODEL CONSTRAINT:
  NEW（indb）；                   ！命名（组间）间接效应
  indb＝a*b;                    ！计算（组间）间接效应
OUTPUT: CINTERVAL;
```

框9-8　基于MSEM的2-1-1模型多水平中介效应分析（随机斜率）

```
TITLE: 2-1-1 mediation（MSEMwith random slope）
DATA: File is 211. dat;
Variable:
  Names are
    schoolid scmatedu readatt readsco greadatt;
  Missing are all（－9999）；
  USEVARIABLES＝scmatedu readatt readsco;
  BETWEEN＝scmatedu;
  CLUSTER＝schoolid;
ANALYSIS:
  TYPE＝TWOLEVEL RANDOM;
MODEL:
  %WITHIN%
```

```
   readsco readatt;                      ! 估计y和m水平-1（残差）协方
差
   s | readsco ON readatt;              ! 组内bw路径作为随机效应进行估
计
  %BETWEEN%
   scmatedu readatt readsco;           ! 估计x、m、y水平-2（残差）协
方差
   readatt ON scmatedu（a）;             ! m对x回归，标记为a
   readsco ON readatt（bb）;              ! y对m回归，标记为bb，为情境
效应，非组间斜率
   readsco ON scmatedu;
   s WITH scmatedu readatt readsco;
  [s]（bw）;                               ! 估计s均值，标记为bw
MODEL CONSTRAINT:
  NEW（b indb）;
  b=bb+bw;                             ! 计算组间b路径
  indb=a*b;                             ! 计算（组间）间接效应
OUTPUT: CINTERVAL;
```

使用R软件利用框9-8模型相关输出结果对中介效应经验分布进行自举法估计，代码示例见框9-9。

框9-9　基于MSEM的2-1-1模型中介效应经验分布R代码

```
#a为水平-2  M对Y的预测效应                #
#astd为a的标准误                #
#bb为水平-2  M对Y的预测效应      #
#bbstd为bb的标准误                #
#bw为M对Y的随机斜率的条件均值      #
#bwstd为bw的标准误                #
#"rep＝20000"定义了重取样的次数为20000    #
#"conf＝95"定义了95%的CI    #

a  ＝0.011
bb＝371.528
bw＝54.589
astd＝0.006
bbstd＝468.818
bwstd＝4.303
rep＝20000
conf＝95
avec  ＝rnorm（rep）*astd＋a
bbvec＝rnorm（rep）*bbstd＋bb
bwvec＝rnorm（rep）*bwstd＋bw
bvec＝bbvec＋bwvec
indb＝avec*bvec
low＝（1－conf/100）/2
upp＝（（1－conf/100）/2）＋（conf/100）
LL＝quantile（indb，low）
UL＝quantile（indb，upp）
```

```
LL4=format (LL, digits=5)
UL4=format (UL, digits=5)
hist ( indb , breaks = 'FD' , col = 'skyblue' , xlab = paste
(conf, '%  Confidence Interval', 'LL', LL4, 'UL', UL4), main
='Distribution of Indirect Effect')
```

表9-4　中介效应2-1-1模型结果

	T89	T91	T93	T97	T00	T04	T06
n	849	761	720	552	571	534	521
均值	88.003	87.837	89.420	92.422	93.988	95.911	96.436
标准差	11.384	12.1786	11.608	13.157	13.133	14.041	14.152
T89	1						
T91	0.527	1					
T93	0.452	0.493	1				
T97	0.467	0.498	0.483	1			
T00	0.357	0.419	0.451	0.556	1		
T04	0.274	0.316	0.296	0.415	0.447	1	
T06	0.319	0.265	0.348	0.406	0.420	0.496	1

从结果来看，2-2-1模型使用单水平模型拟合，间接效应显著，但使用MLM、MSEM后，不再显著。具体来看，是组间b路径不显著。另外，随机斜率s并不显著，实际建模应该考虑固定斜率模型。

（3）1-1-1模型

以PISA数据中水平-1学生家庭社会经济地位hisei（X）、家庭文化资源hedres（M）和阅读得分readsoc（*Y*）3个变量为例，分别在单水平和多水平框架下进行中介效应的分析。由于模型较为复杂，为了

便于迭代收敛，hisei和readsoc除以100处理。相关代码见框9-10至框9-14。

框9-10　1-1-1模型单水平中介效应分析

```
TITLE: 1-1-1 Simple Mediation Analysis;
DATA:   File is 111. dat;
Variable:
  Names are
    schoolid hisei hedres ghedres readsco ghisei;
  Missing are all（－9999）;
  USEVARIABLES＝hedres hisei readsco;
Define: readsco＝readsco／100;
    hisei＝hisei／100;
ANALYSIS:
  BOOTSTRAP IS 5000;
MODEL:                  ! X: hisei Y: readsco M: hedres
  hedres ON hisei;    ! a
  readsco ON hedres; ! b
  readsco ON hisei; ! c'
MODEL INDIRECT:
  readsco IND hedres hisei;
OUTPUT: CINTERVAL（BCBOOTSTRAP）;
```

框9-11　基于MLM的1-1-1模型多水平中介效应分析

```
TITLE: Multilevel Mediation Analysis with random slope, 1-1-1
Model, traditional MLM;
DATA:   File is 111. dat;
Variable:
  Names are
    schoolid hisei hedres ghedres readsco ghisei;
  Missing are all (－9999);
  USEVARIABLES＝hedres hisei readsco;
  WITHIN＝hisei;
  BETWEEN＝;
  CLUSTER＝schoolid;
Define: readsco＝readsco/100;
    hisei＝hisei/100;
ANALYSIS:
  TYPE＝TWOLEVEL RANDOM;
MODEL:
  %WITHIN%
   sa | hedres ON hisei;     ! a
   sb | readsco ON hedres; ! b
   sc | readsco ON hisei; ! c'
  %BETWEEN%
    sa sb sc hedres readsco;                ! 估计水平-2(残差)
方差
    [sa] (a);       ! 估计sa均值,命名为a
    [sb] (b);     ! 估计sb均值,命名为b
   sa WITH sc hedres readsco;
```

```
  sb WITH sc hedres readsco;
  sc WITH hedres readsco;
  readsco WITH hedres;
  sa WITH sb（cab）；                ！估计水平-2  sa与sb协方差，
命名为cab
MODEL CONSTRAINT:
 NEW（ind）；           ！命名间接效应
 ind＝a*b＋cab;         ！计算间接效应
OUTPUT: CINTERVAL;
```

框9-12　基于Unconflated MLM的1-1-1模型多水平中介效应分析

```
TITLE : Multilevel Mediation Analysis , 1-1-1  Model ,
unconflated MLM;
DATA:  File is 111. dat;
Variable:
 Names are
  schoolid hisei hedres ghedres readsco ghisei;
 Missing are all（－9999）；
  USEVARIABLES = schoolid readsco hedres hisei ghisei
ghedres;
 WITHIN＝hisei hedres;
 BETWEEN＝ghisei ghedres;
 CLUSTER＝schoolid;
Define: readsco＝readsco／100;
```

```
   hisei＝hisei／100;
   CENTER hisei hedres  （GROUPMEAN）;       ！hisei hedres组均值中心化
ANALYSIS:
  TYPE＝TWOLEVEL RANDOM;
MODEL:
  %WITHIN%
   hedres ON hisei（aw）;          ！aw
   readsco ON hedres（bw）;   ！bw
   readsco ON hisei;     ！cw'
   [hedres@0];          ！hedres已被组中心化，所以均值固定为零
  %BETWEEN%
   ghedres readsco;
   ghedres ON ghisei（ab）;
   readsco ON ghedres（bb）;
   readsco ON ghisei;
MODEL CONSTRAINT:
  NEW（indb indw）;              ！命名间接效应
  indw＝aw*bw;   ！计算组内间接效应
  indb＝ab*bb;   ！计算组间间接效应
OUTPUT: CINTERVAL;
```

框9-13　基于MSEM的1-1-1模型多水平中介效应分析（固定斜率）

```
TITLE: Multilevel Mediation Analysis, 1-1-1 Model, with fixed
slopes, MSEM;
DATA:   File is 111. dat;
Variable:
  Names are
    schoolid hisei hedres ghedres readsco ghisei;
  Missing are all (－9999) ;
  USEVARIABLES＝hedres hisei readsco;
  CLUSTER＝schoolid;
Define: readsco＝readsco／100;
    hisei＝hisei／100;
ANALYSIS:
  TYPE＝TWOLEVEL;
MODEL:
  %WITHIN%
   hedres ON hisei (aw) ;        ! aw
   readsco ON hedres (bw) ; ! bw
   readsco ON hisei;    ! cw'
  %BETWEEN%
   hisei hedres readsco;        ! 估计水平-2(残差)方差
   hedres ON hisei (ab) ;
   readsco ON hedres (bb) ;
   readsco ON hisei;
MODEL CONSTRAINT:
  NEW (indb indw) ;              ! 命名间接效应
  indw＝aw*bw;             ! 计算组内间接效应
```

```
  indb＝ab*bb;              ！计算组间间接效应
OUTPUT: CINTERVAL;
```

框9-14 基于MSEM的1-1-1模型多水平中介效应分析（随机斜率）

```
TITLE： Multilevel Mediation Analysis， 1-1-1  Model， with
Random slopes，MSEM;
DATA:  File is 111. dat;
Variable:
  Names are
    schoolid hisei hedres ghedres readsco ghisei;
  Missing are all（－9999）;
  USEVARIABLES＝hedres hisei readsco;
  CLUSTER＝schoolid;
Define: readsco＝readsco／100;
    hisei＝hisei／100;
ANALYSIS:
  TYPE＝TWOLEVEL RANDOM;
MODEL:
  ％WITHIN％
   sa｜hedres ON hisei;                  ！组内aw路径作为随机效应
进行估计
   sb｜readsco ON hedres;               ！组内bw路径作为随机效应进行
估计
   sc｜readsco ON hisei;
```

```
  %BETWEEN%
   sa sb sc hisei hedres readsco;      ！估计水平-2（残差）方差
   sa WITH sc hisei hedres readsco;
   sa WITH sb（cab）;         ！估计水平-2  sa与sb协方差，命名为cab
   sb WITH sc hisei hedres readsco;
   sc WITH hisei hedres readsco;
   hedres ON hisei（ab）;      ！ab，注意这里为情境效应，非组间斜率
   readsco ON hedres（bb）;      ！bb，注意这里为情境效应，非组间斜率
   readsco ON hisei;
    [sa]（aw）;      ！估计sa均值，命名为aw
    [sb]（bw）;      ！估计sb均值，命名为bw
MODEL CONSTRAINT:
NEW（a b indb indw）;                 ！命名间接效应
  a=aw+ab;              ！计算组间a路径
  b=bw+bb;              ！计算组间b路径
  indw=aw*bw+cab;       ！计算组内间接效应
  indb=ab*bb;           ！计算组间间接效应
OUTPUT: CINTERVAL;
```

以R软件利用框9-14模型相关输出结果对中介效应经验分布进行自举法估计，代码示例见框9-15。

框9-15　基于MSEM的1-1-1模型中介效应经验分布R代码

```
#aw为X对M随机斜率（sa）条件均值                    #
#awstd为aw的标准误            #
#ab为水平-2  X对M的预测效应          #
#abstd为ab的标准误          #
#bw为M对Y随机斜率（sb）条件均值       #
#bwstd为bw的标准误          #
#bb为水平-2  M对Y的预测效应       #
#bbstd为bb的标准误          #
#cab为sa和sb协方差          #
#cabstd为cab的标准误      #
#"rep=20000"定义了重取样的次数为20000          #
#"conf=95"定义了95%的CI         #

aw  =0.921
ab=3.889
bw=0.116
bb=0.654
cab=0.007
awstd=0.083
abstd=0.868
bwstd=0.015
bbstd=0.643
cabstd=0.018
rep=20000
conf=95
awvec=rnorm（rep）*awstd+aw
```

```
abvec=rnorm（rep）*abstd+ab
bwvec=rnorm（rep）*bwstd+bw
bbvec=rnorm（rep）*bbstd+bb
cabvec=rnorm（rep）*cabstd+cab
avec=awvec+abvec*bvec=bwvec+bbvec
indw=awvec*bwvec+cabvec
indb=abvec*bbvec
low=（1－conf／100）／2
upp=（（1－conf／100）／2）+（conf／100）
LLw=quantile（indw，low）
ULw=quantile（indw，upp）
LLw4=format（LL，digits=5）
ULw4=format（UL，digits=5）
hist （ indw ， breaks = 'FD' ， col = 'skyblue' ， xlab = paste
（conf，'%Confidence Interval'，'LLw'，
LLw4，'ULw'，ULw4），main='Distribution of Indirect Effect'）
LLb=quantile（indb，low）
ULb=quantile（indb，upp）
LLb4=format（LLb，digits=5）
ULb4=format（ULb，digits=5）
hist （ indb ， breaks = 'FD' ， col = 'skyblue' ， xlab = paste
（conf，'%Confidence Interval'，'LLb'，LLb4，'ULb'，ULb4），
main='Distribution of Indirect Effect'）
```

从结果来看，区分组间和组内效应后，组内间接效应统计上显著。具体来看，在随机斜率模型中，组间b路径、bb（情境效应），统

计上不显著。意味着家庭文化资源对阅读得分在学校层面没有显著关系。结合个体层面结果可知，间接作用主要体现在个体层面。从相关输出参数来看，随机斜率模型拟合得要比固定斜率模型好，不过sa、sb方差统计上并不显著，实际建模时可考虑进一步对模型进行修订。

表9-5　中介效应1-1-1模型结果

模　型	位　置	a	b	间接效应[95%置信区间]
单水平	未区分	1.325(0.078)	0.181(0.015)	0.240[0.197, 0.291]
traditional MLM	混淆	1.009(0.080)	0.116(0.015)	0.120[0.071, 0.169]
unconflated MLM	组内	0.933(0.084)	0.122(0.015)	0.114[0.079, 0.149]
	组间	0.031(0.002)	0.345(0.091)	0.011[0.005, 0.017]
MSEM，固定斜率	组内	0.932(0.084)	0.122(0.015)	0.114[0.079, 0.148]
	组间	3.567(0.283)	0.275(0.210)	0.981[−0.488, 2.449]
MSEM，随机斜率	组内	0.921(0.083)	0.116(0.015)	0.114[0.064, 0.163]
	组间	3.889(0.868)	0.654(0.643)	2.544[−2.618, 7.706]

9.4.2　高水平结果模型

相比于MLM，MSEM处理多水平中介效应的一个优势是可以处理结果变量处于高层的情况。以下代码为高水平结果的示例。

（1）2-1-2模型

以水平-2教师参与tchparti（X）、水平-1师生关系评价studrel（M）和水平-2学生行为studbeha（*Y*）为例，构建2-1-2多水平中介效应模型。代码见框9-16。

框9-16　基于MSEM的2-1-2模型多水平中介效应分析

```
TITLE: Multilevel Mediation Analysis，2-1-2 Model，MSEM;
DATA: FILE IS 212. dat;
VARIABLE:
  NAMES ARE schoolid studrel studbeha tchparti;
  USEVARIABLES ARE schoolid studrel studbeha tchparti;
  Missing are all（－9999）;
  CLUSTER IS schoolid;
  BETWEEN ARE studbeha tchparti;
ANALYSIS:
  TYPE IS TWOLEVEL;
MODEL:
  %WITHIN%
    studrel;                        ！估计m水平-2（残差）方差
  %BETWEEN%
    studbeha tchparti;              ！估计x、y水平-2（残差）方差
    studrel ON tchparti（a）;       ！m对x回归，斜率标记为a
    studbeha ON studrel（b）;       ！y对m回归，斜率标记为b
    studbeha ON tchparti;
MODEL CONSTRAINT:
    NEW（indb）;                        ！命名间接效应
    Indb＝a*b;                          ！计算组间间接效应
OUTPUT: CINTERVAL;
```

（2）1-2-1模型

以水平-1学生家庭社会经济地位hisei（X）、水平2学生行为studbeha（M）和水平-1学生阅读成绩readsoc（Y ）为例，构建1-2-1多水平中介效应模型。代码见框9-17。

框9-17　基于MSEM的1-2-1模型多水平中介效应分析

```
TITLE: Multilevel Mediation Analysis，1-2-1 Model，MSEM;
DATA:
  FILE IS 121. dat;
VARIABLE:
  NAMES ARE schoolid hisei studbeha readsco;
  USEVARIABLES schoolid hisei readsco studbeha;
  Missing are all（－9999）;
  CLUSTER IS schoolid;
  BETWEEN ARE studbeha;
ANALYSIS:
  TYPE IS TWOLEVEL RANDOM;
MODEL:
  %WITHIN%
    readsco ON hisei;
  %BETWEEN%
    hisei studbeha readsco;          ！估计x、m、y水平-2（残差）方
差
    studbeha ON hisei（a）;     ！m对x回归，标记为a
    readsco ON studbeha（b）;     ！y对m回归，标记为b
    readsco ON hisei;
```

```
MODEL CONSTRAINT：
  NEW（indb）；                           ！命名间接效应
  indb＝a*b；                             ！计算组间间接效应
OUTPUT：CINTERVAL；
```

（3）1-2-2模型

以水平-1学生家庭社会经济地位hisei（X）、水平-2学生行为studbeha（M）和阅读平均得分greadsco（*Y*）为例，构建1-2-2多水平中介效应模型。代码见框9-18。

框9-18　基于MSEM的1-2-2模型多水平中介效应分析

```
TITLE：Multilevel Mediation Analysis，1-2-2 Model，MSEM；
DATA：
  FILE IS 122. dat；
VARIABLE：
  NAMES ARE schoolid hisei studbeha greadsco；
  USEVARIABLES ARE schoolid hisei studbeha greadsco；
  Missing are all（－9999）；
  CLUSTER IS schoolid；
  BETWEEN ARE studbeha greadsco；
ANALYSIS：TYPE IS TWOLEVEL；
MODEL：
```

```
  %WITHIN%
   hisei;
  %BETWEEN%
   studbeha greadsco;        ！估计m、y水平-2（残差）方差
   studbeha ON hisei（a）；      ！m对x回归，斜率标记为a
   greadsco ON studbeha（b）；！y对m回归，斜率标记为b
   greadsco ON hisei;
MODEL CONSTRAINT:
  NEW（indb）；                              ！命名间接效应
  indb＝a*b;                                ！计算组间间接效应
OUTPUT: CINTERVAL;
```

（4）1-1-2模型

以水平-1学生家庭社会经济地位hisei（X）、阅读成绩readsco（M）和水平-2学生行为studbeha（*Y*）为例，构建1-1-2多水平中介效应模型。代码见框9-19。

框9-19　基于MSEM的1-1-2模型多水平中介效应分析

```
TITLE: Multilevel Mediation Analysis，1-1-2 Model，MSEM;
DATA: FILE IS 112. dat;
VARIABLE:
  Names are
```

```
    schoolid hisei studbeha readsco;
  USEVARIABLES ARE hisei studbeha readsco;
  Missing are all (－9999);
  CLUSTER IS schoolid;
  BETWEEN ARE studbeha;
ANALYSIS:
    TYPE IS TWOLEVEL RANDOM;
MODEL:
  %WITHIN%
   readsco ON hisei;
  %BETWEEN%
   readsco studbeha;           ! 估计m、y水平-2（残差）方差
   readsco ON hisei (a);        ! m对x回归，斜率标记为a
   studbeha ON readsco (b); ! y对m回归，斜率标记为b
   studbeha ON hisei;
MODEL CONSTRAINT:
  NEW (indb);                              ! 命名间接效应
  indb＝a*b;                              ! 计算组间间接效应
OUTPUT: CINTERVAL;
```

如前所述，包含水平-2变量中介效应，必定是组间中介效应。而且，除了1-1-1模型和带有随机斜率的2-1-1模型，这里考虑的其他模型的中介效应也只存在于组间水平。从代码书写规范来讲，基于MSEM构建多水平中介效应分析构建模型的基本原则是：对于模型中涉及中介效应的变量，水平-2测量的变量只在水平-2出现，水平-1测量的变量（除非严格界定只属于水平-1，无水平-2显著变化）同时在水平-1

和水平-2出现，然后各自在不同水平构建相应方程。由此，基于MSEM方式，处于不同水平的变量之间的中介效应关系，被区分为组间和组内成分后在各自所属的水平进行处理，模型层次清楚，代码书写较之于基于MLM的多水平中介效应方式也大为简化。

9.4.3 其他拓展

（1）纳入潜变量

多水平中介效应模型可以方便地拓展到含有潜变量的情况。以1-1-2模型为例，假定需要研究水平-1家庭文化资本（X：cultposs、hisei、hedres）、学生成绩（M：readsco mathsco sciesco）和水平-2学生表现（Y：sc17q02、sc17q03、sc17q04、sc17q05）之间的多水平中介效应，模型代码见框9-20。

框9-20 基于MSEM的1-1-2模型中介效应分析（含潜变量）

```
TITLE: Multilevel Mediation Analysis, 1-1-2 Model,
      with Latent Variable; MSEM;
DATA:   FILE IS 112_latent. dat;
VARIABLE:
  NAMES ARE
     schoolid hisei cultposs hedres sc17q02 sc17q04 sc17q07
sc17q08 sc17q10
    sc17q12 readsco mathsco sciesco;
```

```
  Missing are all (－9999);
  USEVARIABLES ARE
     hisei  cultposs  hedres  sc17q02  sc17q04  sc17q07  sc17q08
sc17q10
    sc17q12 readsco mathsco sciesco;
  Between＝sc17q02 sc17q04 sc17q07 sc17q08 sc17q10 sc17q12;
  CLUSTER IS schoolid;
ANALYSIS:
  TYPE IS TWOLEVEL;
  ESTIMATOR＝MLR;
MODEL:
  %WITHIN%
   mw BY readsco mathsco sciesco;
   xw BY cultposs hisei hedres;
   mw ON xw;
  %BETWEEN%
   mb BY readsco mathsco sciesco;
   xb BY cultposs hisei hedres;
   yb BY sc17q02 sc17q04 sc17q07 sc17q08 sc17q10 sc17q12;
   yb ON mb (b);
   mb ON xb (a);
   yb ON xb;
MODEL CONSTRAINT:
  NEW (ab);
  ab＝a*b;
OUTPUT: CINTERVAL;
```

（2）1-（1，1）-1模型

基于MSEM的多水平中介分析方法也可以推广到更复杂的多水平中介模型中。比如，可在1-1-1模型基础上增加中介变量，进行多中介变量间接效应的分析。以1-（1，1）-1模型为例，分析水平-1家庭藏书量books（M1）、阅读态度readatt（M2）家庭社会经济地位hisei（X）、阅读得分pv1read（*Y*）。代码见框9-21。

框9-21　基于MSEM的1-（1，1）-1模型多水平中介效应分析

```
TITLE: Multilevel Mediation Analysis，1-（1，1）-1  Model，MSEM
DATA: FILE IS 1111. dat;
VARIABLE:
  NAMES ARE
   schoolid books hisei pv1read readatt;
  MISSING ARE ALL（－9999）;
  USEVARIABLES ARE
   schoolid books hisei pv1read readatt;
  CLUSTER IS schoolid;
ANALYSIS:
  TYPE IS TWOLEVEL RANDOM;
MODEL:
  %WITHIN%
      pv1read ON books（bw1）;           ! y对m1回归，标记斜率为
bw1
```

```
      pv1read ON readatt（bw2）;          ! y对m2回归，标记斜率为bw2
      s|pv1read ON hisei;          ! y对x回归，标记随机斜率为s
      books WITH readatt;          ! 估计水平-1m1、m2（残差）协方差
      books ON hisei（aw1）;        ! m1对x回归，标记斜率为aw1
      readatt ON hisei（aw2）;        ! m2对x回归，标记斜率为aw2
  %BETWEEN%
       s books readatt pv1read;          ! 估计s、m1、m2和y水平-2（残差）方差
      s WITH books readatt pv1read;         ! 估计s和m1、m2和y的（残差）协方差
      pv1read ON books（bb1）;        ! y对m1回归，标记斜率为bb1
      pv1read ON readatt（bb2）;        ! y对m2回归，标记斜率为bb2
      pv1read ON hisei;        ! y对x回，情境效应，非组间斜率
      books WITH readatt;        ! 估计m1和m2残差协方差
      books ON hisei（ab1）;            ! m1对x回归，标记斜率为ab1
      readatt ON hisei（ab2）;        ! m2对x回归，标记斜率为ab2
        [s];          ! 估计s的均值
MODEL CONSTRAINT:
  NEW（abw1  abw2  abb1  abb2  conw conb）;
  abw1＝aw1*bw1;              ! 计算组内间接效应1
  abw2＝aw2*bw2;         ! 计算组内间接效应2
  abb1＝ab1*bb1;         ! 计算组间间接效应1
  abb2＝ab2*bb2;         ! 计算组间间接效应2
  conw＝abw1-abw2;        ! 组内间接效应比较
```

```
  conb＝abb1-abb2；        ！组间间接效应比较
OUTPUT：CINTERVAL；
```

相较基于MLM方法，基于MSEM的多水平中介效应分析更具弹性，可解决一些复杂的问题，在概念和操作上均显现出优势，但也存在一定局限。比如，基于MSEM的中介效应分析对样本量的要求更高，基于MLM多水平中介效应分析适合小样本的情况。Li＆Beretvas（2013）的模拟研究表明，对于2-2-1模型，尽管MSEM法可以更好地估计中介效应大小，但组群样本量小于80时，遇到了严重的迭代不收敛的问题，因此建议只有组群样本量大于80，并且是最简单的多水平中介效应分析才可以使用基于MSEM的方法。相比之下，基于MLM的方法对样本量要求较低。McNeish（2017）的模拟研究发现，仅10个组群即可满足需求。此外，由于需要的运算量相对较大，如果模型复杂，涉及变量较多，基于MSEM的多水平中介效应分析更容易出现迭代不收敛等情况，尤其是组群样本量小、水平-1变量ICC较低的情况下。具体而言，在情况良好（比如间接效应值较大，组内样本量较大）时，基于MSEM的多水平中介效应分析水平-2推荐样本量为50；而基于MLM，组群样本量15～30即够。

鉴于MSEM估计多水平中介效应得到准确的中介效应以统计功效的降低、需要大样本（尤其水平-2）为代价，有研究者（方杰等，2014）建议可以采取Ledgerwood＆Shrout（2011）针对单水平中介效应分析的“两步法”的策略：在进行多水平中介分析时，中介效应的点估计值以MSEM的分析结果为准，中介效应的显著性以MLM的分析结果为准。

9.5　与调节效应结合的多水平中介效应分析

目前可以在MLM框架下进行有中介的调节效应或者有调节的中介效应（参见Bauer et al，2006）的建模分析，但相比MSEM的策略，可能存在偏差（Preacher et al，2010）。刘东等（2012）提供了基于多水平路径分析的有中介的调节效应或者有调节的中介效应示例。然而，从MSEM框架来看，还不够彻底。因为模型中只有水平-1的内生变量才会出现在水平-2（实际上相当于MLM中的结果变量的处理），而非使用多水平潜协变量（MLC）处理方式将各预测变量（如果可能且必要）同时进行组内／组间的分解。

以下按照刘东等（2012）的思路，在一般多水平路径分析的框架下给出若干常见多水平框架下中介效应和调节效应分析相结合的实例，以供参考。掌握基本的建模原则、策略和代码书写规范后，其他类似模型可由此进一步拓展。

9.5.1　多水平有中介的调节效应分析

预测变量和调节变量通过交互作用影响中介变量，中介变量进而对结果变量产生影响，即所谓有中介的调节效应（Baron＆Kenny，1986；Edwards＆Lambert，2007）。另外一种不同类型的被中介的调节作用是，中介变量针对调节变量对预测变量和结果变量之间关系的调节作用起中介作用（Grant＆Berry，2011）。换言之，中介变量替

代了原来的调节变量，改变了预测变量和结果变量之间的关系。[4]以下将第一种情况称为第Ⅰ种类型的被中介的调节效应，第二种情况属于类型Ⅱ被中介的调节效应。

（1）两水平类型Ⅰ有中介的调节作用

如果研究者在分析中排除任何自下而上的关系（低层次变量影响高层次变量），那么多（两）水平类型Ⅰ被中介的调节作用有3种类型（刘东等，2012）：第一种类型为处于水平-1的中介变量对两个水平-2变量之间交互的中介作用，这个中介变量在水平-1进行测量，但是同时在水平-2也具备组体间方差。严格来讲，这情况下的“真实”中介变量其实是该中介变量的组间方差部分。第二种类型与第一种类型相似，只是中介变量是在水平-2进行测量。在第三种类型中，所有变量都是在水平-1进行测量，但是又都在水平-2具备组间方差。具体研究时可在群体内部检验类型Ⅰ被中介的调节作用（使用组中心化值），同时允许变量在水平-2部分彼此自由相关。

以水平-2学校教育资源质量scmatedu（*X*）、教师参与tchparti（*W*），水平-1总结元认知metasum（*M*）和阅读得分readsco（*Y*）为例，构建两水平类型Ⅰ有中介的调节作用路径分析模型。概念图如图9-9所示。该实例属于上述第一种情况。

实际建模时，可遵照图9-10所示进行模型设定。从图9-10可以看到，metasum（*M*）和阅读得分readsco（*Y*）作为内生变量，属于组内/组间两用，有组间方差。水平-2基本上接近于单水平的有中介的

调节效应模型路径图（参见王孟成，2014：49），只是针对的M、Y 的组间部分。

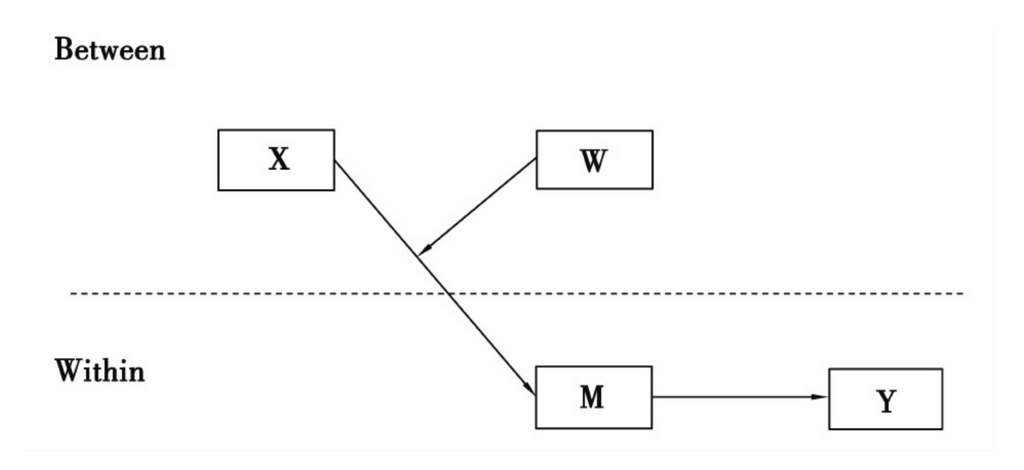


图9-9　两水平类型Ⅰ有中介的调节作用概念图

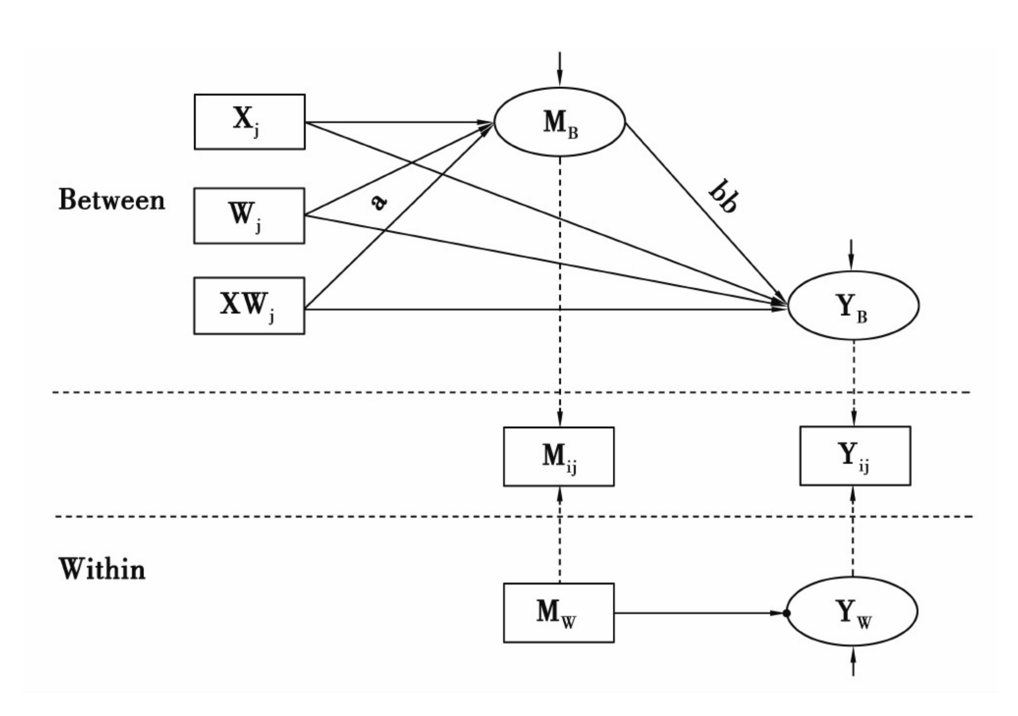


图9-10　两水平类型Ⅰ有中介的调节作用模型图

如前所述，X、M、Y 中只要有一个变量属于水平-2，则中介效应必发生于组间。在此固定斜率的两水平类型Ⅰ有中介的调节效应模型中，中介效应体现于水平-2，属于组间中介效应模型。在有随机斜率

模型中，需要考虑水平-1中$M \to Y$ 的组内路径系数。相应的代码见框9-22。[(5)]

框9-22　两水平类型Ⅰ有中介的调节作用

```
Title: A two-level Type I mediated moderation path analysis model
    Scmatedu（X）and tchparti（W）are level 2,
    metasum（M）and readsco（Y

）are level 1，and have between－group variance;
DATA:
    FILE IS 2-211. dat;
VARIABLE:
    NAMES ARE schoolid metasum scmatedu tchparti readsco;
       USEVARIABLES ARE scmatedu metasum tchparti readsco scmatedutchparti;
    Missing are all（－9999）;
    CLUSTER＝schoolid;
    WITHIN＝;
    BETWEEN＝scmatedu tchparti scmatedutchparti;
DEFINE:
    CENTER scmatedu tchparti（GRANDMEAN）;  ！水平-2预测变量、调节变量总中心化
    scmatedutchparti＝scmatedu*tchparti;
ANALYSIS: TYPE＝TWOLEVEL;
```

```
MODEL：
    %WITHIN%
      readsco on metasum（bw）    ！y与m组内关系，当前模型中与效应量计算无关
    %BETWEEN%
              metasum on scmatedu tchparti
                    scmatedutchparti  （a）；              ！被中介的交互效应，标记为a
    readsco on metasum  （bb）                     ！y与m组间关系，标记为bb
                  scmatedu tchparti scmatedutchparti；    ！控制其他变量
MODEL CONSTRAINT：
    NEW（ind）；
    ind＝a*bb;
OUTPUT：SAMPSTAT CINTERVAL;
```

（2）两水平类型Ⅱ有中介的调节作用

类型Ⅱ有中介的调节效应有两种可能形式（刘东等，2012）：调节变量（W）、中介变量（M）处于水平-2，预测变量（X）、结果变量（Y）处于水平-1；调节变量（W）处于水平-2，预测变量（X）、中介变量（M）和结果变量（Y）处于水平-1，如果预测变量（X）、调节变量（W）、中介变量（M）均处于水平-2，但结果变量处于水平-1。该模型事实上预测的是水平-1变量的组间方差成分，其实

可以看作水平-2的单层次研究。同样，如果四个变量均处于水平-1，也可以视为单水平的研究，同时允许变量组间方差自由相关。

以下为第一种两水平类型Ⅱ有中介的调节效应模型实例。教师参与tchparti（*W*）和学生行为studbeha（*M*）属于水平-2变量，学生家庭地位hisei（*X*）属于水平-1变量，阅读成绩readsco（*Y*）水平1测量，但作为结果变量，存在组间方差，可视为组间/组内两用变量。该模型实际上可以看作把W对X→Y的调节作用分为两部：一部分是直接的调节作用，另一部分是通过中介变量M的调节作用。

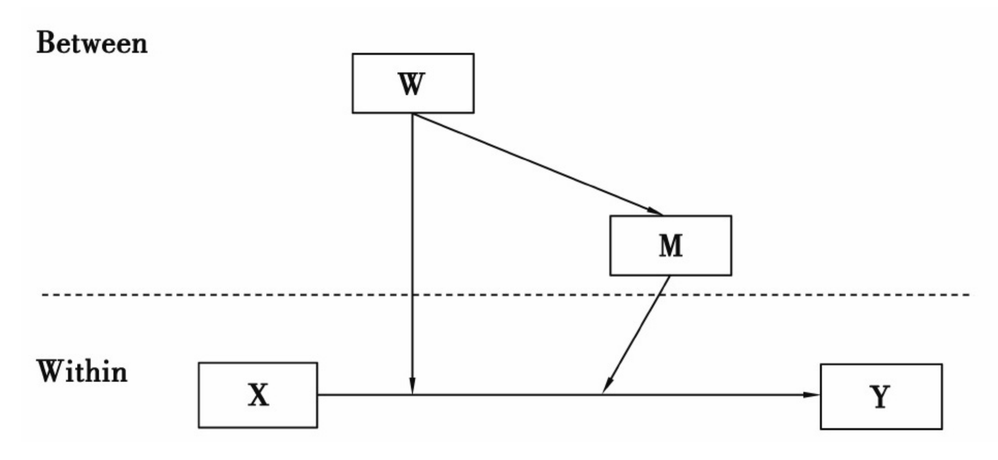


图9-11　两水平类型Ⅱ有中介的调节作用概念图

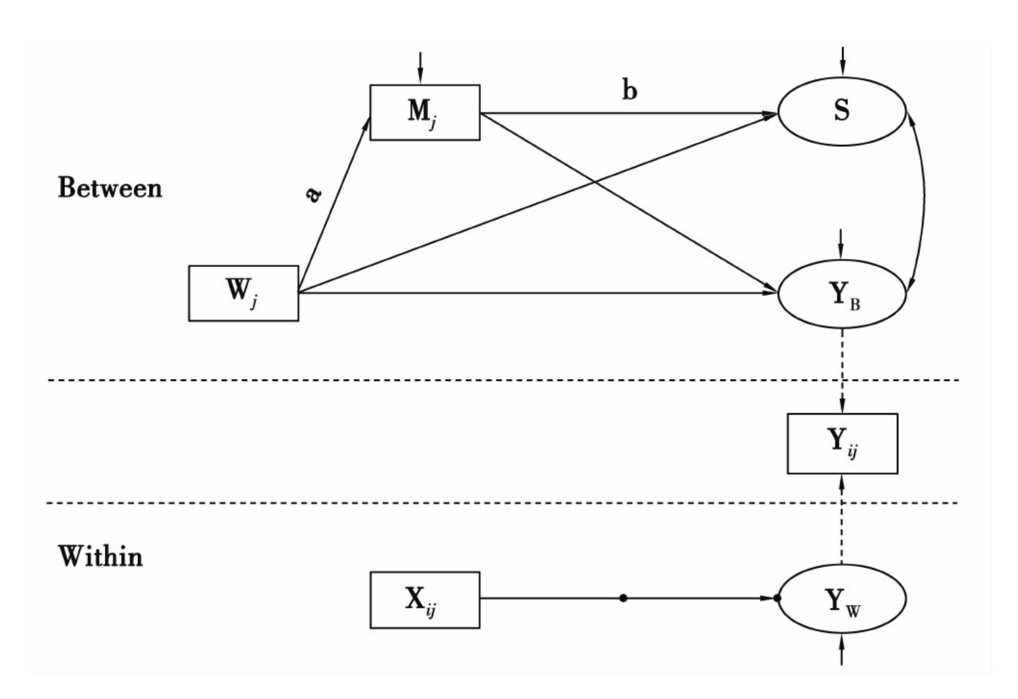


图9-12　两水平类型Ⅱ有中介的调节作用模型图

代码见框9-23，注意因为预测变量（X）和调节变量（W）隶属于不同的层面，不能采取直接生成交互项的形式，但可以利用随机斜率的策略生成跨层交互作用，对水平-1变量产生作用。该模型的中介、调节作用实际上发生于组间，处于水平-2的$X \rightarrow Y$的随机斜率（均值）实际上可视为水平-1层面X对Y效应的体现。因此从模型来看，可以在水平-2直接针对s进行设定。W对$X \rightarrow Y$的调节效应的两个组成部分分别对应WX和MX两个跨层交互项，后者属于被中介的调节效应。

框9-23　两水平类型Ⅱ有中介的调节作用

```
TITLE： A two-level Type Ⅱ mediated moderation path analysis model，
     tchparti（w）and studbeha（m）are level 2，hisei（x）and readsco（y）are level  1;
```

```
DATA: FILE IS 2-121. dat;
VARIABLE:
    NAMES ARE schoolid hisei studbeha tchparti readsco;
    USEVARIABLES ARE hisei studbeha tchparti readsco;
    CLUSTER＝schoolid;
    WITHIN＝hisei;
    BETWEEN＝tchparti studbeha;
DEFINE:
       CENTER tchparti studbeha（GRANDMEAN）;    ！水平-2调节变量、中介变量总中心化
      CENTER hisei（GROUPMEAN）;      ！水平-1预测变量组中心化
ANALYSIS: TYPE＝TWOLEVEL RANDOM;
MODEL:
    %WITHIN%
      s | readsco on hisei;
    %BETWEEN%
          studbeha on tchparti（a）;    ！w与m之间的效应，标记为a
          s on studbeha（b）     ！s对m的回归，会自动形成组内x与m的交互项，标记为b
                    tchparti;
      readsco on studbeha tchparti;
      readsco with s;
MODEL CONSTRAINT:
    NEW（ind）;
    ind＝a*b;
OUTPUT: SAMPSTAT CINTERVAL;
```

9.5.2　多水平有调节的中介效应分析

以下为两水平第一阶段有调节的中介效应模型实例。其中，学生行为studbeha（W）是调节变量，处于水平-2；学生家庭社会经济地位hisei（X）、家庭文教育资源hedres（M）、阅读得分readsco（Y）处于水平-1，其中hedres（M）和readsco（Y）有组间方差变化，属于组间/组内两用变量。studbeha（W）对$X \rightarrow M \rightarrow Y$中介效应第一阶段有调节作用。

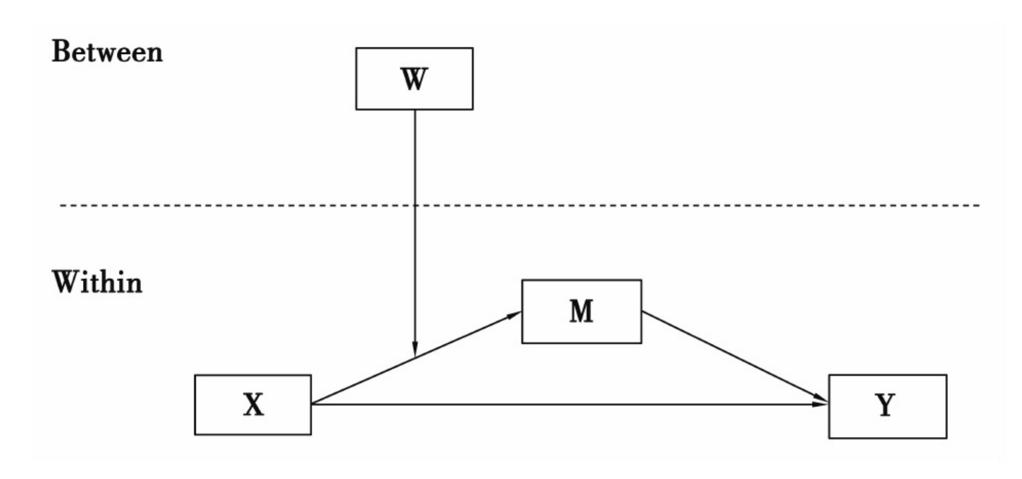


图9-13　两水平第一阶段有调节的中介效应概念图

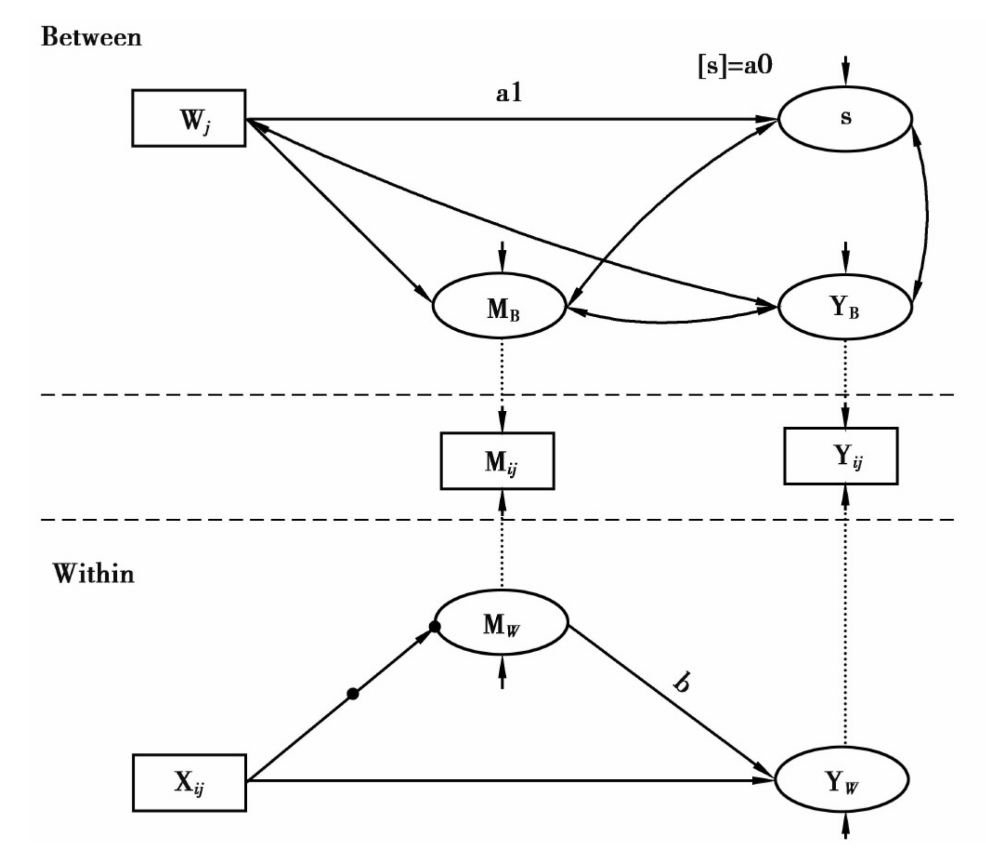


图9-14　两水平第一阶段有调节的中介效应模型图

由于调节变量W处于水平-2，而预测变量X处于水平-1，无组间变化，因此该模型的调节作用需要使用随机斜率的方法来实现。因为是第一阶段的调节，WX交互项作用于M的组内部分，其中包括了固定的跨层交互作用a 1（即XW的系数）以及随机变化的部分，可通过对s取均值估计（a 0）。b路径存在于组内。

相应代码见框9-24。其中，为了减轻因变量测度尺度悬殊造成的模型迭代不易收敛的问题，连续变量hisei和readsco均除以100。单位的变化不改变变量之间的实质关系，只是结果解释的时候注意一下即可。

框9-24　两水平第一阶段有调节的中介效应模型

```
TITLE : A two-level first-stage moderated mediation path analysis model,
      studbeha (w) is level 2, hisei (x), hedres (m), readsco (y) are level 1,
      hedres (m) and readsco (y) both have between-group variance;
DATA: FILE IS 2-111. dat;
VARIABLE:
  NAMES ARE schoolid hisei hedres studbeha readsco;
  USEVARIABLES ARE hisei hedres studbeha readsco;
  Cluster＝schoolid;
  WITHIN＝hisei;
  BETWEEN＝studbeha;
DEFINE:
  readsco＝readsco/100;
  hisei＝hisei/100;
  CENTER studbeha (GRANDMEAN); ! 水平-2调节变量w总中心化
  CENTER hisei (GROUPMEAN);     ! 水平-1预测变量x组中心化
ANALYSIS: TYPE＝TWOLEVEL RANDOM;
MODEL:
  %WITHIN%
    s|hedres on hisei;
    readsco on hedres (b)
          hisei;
  %BETWEEN%
    s on studbeha (a1); ! 跨层交互作用
```

```
    [s]（a0）;          ! 随机斜率均值
    hedres on studbeha;
    hedres with s;
    readsco with hedres;
    readsco with s;
    readsco with studbeha;
MODEL CONSTRAINT:
  NEW（ind_h ind_l diff）;
     ind_h=（a0+a1*（.8035087））*b;      ! w均值为0，SD为.8035087
    ind_l=（a0+a1*（-.8035087））*b;
    Diff=ind_h-ind_l;                 ! 显著表明有调节的中介效应存在
OUTPUT: SAMPSTAT CINTERVAL;
```

以R软件利用框9-24模型相关输出结果对上述两水平第一阶段有调节的中介效应经验分布进行自举法估计，代码示例见框9-25。

框9-25　两水平有调节的中介效应经验分布R代码

```
#a0为X对M随机斜率的条件均值                #
#a0std为a0的标准误                        #
#a1为W对X与M之间随机斜率的预测效应          #
#a1std为a1的标准误              #
#b为M对Y的固定效应            #
```

```
#bstd为b的标准误              #
#"rep=20000"定义了重取样的次数为20000#
#"conf=95"定义了95%的CI           #
#取W的+/-1  SD的变化，SD=0.8035087        #

a0=0.931
a1=-0.153
b=0.122
a0std=0.082
a1std=0.115
bstd=0.015
rep=20000
conf=95
a0vec=rnorm（rep）*a0std+a0
a1vec=rnorm（rep）*a1std+a1
bvec=rnorm（rep）*bstd+b
amhvec=a1vec*（0.8035087）+a0vec
amlvec=a1vec*（-0.8035087）+a0vec
abh=amhvec*bvec
abl=amlvec*bvec
d=abh-abl
low=（1-conf/100）/2
upp=（（1-conf/100）/2）+（conf/100）
LL=quantile（d，low）
UL=quantile（d，upp）
LL4=format（LL，digits=5）
UL4=format（UL，digits=5）
```

```
hist （ d ， breaks ＝ 'FD' ， col ＝ 'skyblue' ， xlab ＝
paste（conf，'％Confidence Interval'，'LL'，LL4，'UL'，UL4），
main＝'Distribution of Indirect Effect'）
```

上述实例为（1）调节变量*W*处于水平-2，预测变量（*X*）、中介变量（*M*）和结果变量（*Y*）均处于水平-1的情况。除此之外，还可能存在（2）调节变量（*W*）处于水平-2，预测变量（*X*）或预测变量（*X*）与中介变量（*M*）处于水平-2，结果变量（*Y*）处于水平-1的情况；以及（3）两个调节变量（*W*1、*W*2）处于不同层次的情况。因为预测变量（*X*）、中介变量（*M*）和结果变量（*Y*）中有变量属于水平-2，后两种情况中介关系仅存在于水平-2（Zhang et al，2009；Preacher et al，2010）。因此，在计算多水平有调节的中介作用时，调节变量在不同取值的间接效应都是在水平-2进行计算。当调节变量在水平-1进行测量时（如第3种情况），该调节变量的组间方差，即其在水平-2的潜均值才是要用到的真正调节变量（刘东等，2012）。

Card（2012）在指出基于MSEM的多水平中介效应分析的局限时，提到基于MSEM的调节效应尚无完整的解决方案。不过，Preacher et al（2016）在最新的一篇论文中对基于MSEM的调节效应分析进行了探讨。与基于MSEM的多水平中介效应分析逻辑一致，其思路是依照MLC的思路，将所有参与建模的非专用变量分解为组间或组内不同部分作为潜变量，然后连同各水平专用变量，以及两用变量分解后的不同部分之间，再根据相关假设通过随机斜率（RCP）或潜交互（LMS）的方法进行跨水平或同水平的调节效应分析。比如，对于一个2×（1→1）的多水平调节效应分析模型（调节变量处于水平-2，预测变量、结果变

量处于水平-1），基于上述方法，可以单独分析Z $_j$ 调节了X_{ij} 对Y_{ij} 的组内（W ）效应，或者Z $_j$ 调节了X_{ij} 对Y_{ij} 的组间（B ）效应。然而，由于此类模型本身往往同时涉及多水平、随机斜率、潜交互作用等情况，运算量巨大且迭代不易收敛，一般需要指定开始值（starting values）才可能得到估计的结果。目前在这样的框架下结合中介效应分析理论上可行，但实际估计起来相当困难。

9.6　本章小结

本章介绍了在多水平模型框架下进行中介效应检验的问题，并对MLM和MSEM两个思路进行了比较。

由于采取了针对多水平潜协变量“分解优先”的思路，基于MSEM的多水平中介效应更富有弹性，不会出现组内/组间效应的混淆，结果也更具信度。因此，若样本条件满足，可优先基于MSEM进行多水平中介效应分析。

结合实例，本章介绍了常见的2-2-1、2-1-1、1-1-1模型和高水平结果模型，以及基于多水平路径分析的与调节效应相结合的多水平中介效应分析。基于MSEM的调节效应分析目前属于前沿领域，在实际估计时仍面临一定困难，有这方面需要的读者可参考“推荐阅读”中提到的相关文献。

※推荐阅读※

方法方面，基于MLM的多水平中介效应分析可参见Krull＆MacKinnon（1999；2001）、Bauer et al（2006）、MacKinnon（2008：237-274）、Tofighi（2010）等文献。

基于MSEM的多水平中介效应分析目前主要是Preacher、Zhang、Zyphur等学者的研究。相关文献可参考Zhang et al（2009）、Preacher et al（2010）、Preacher et al（2011）、刘东等（2012）等文献。Pituch＆Stapleton（2011）对基于MLM和SEM多水平中介效应进行了对比性的介绍。中文文献中，温忠麟等（2012）、方杰等（2014）对多水平中介分析进行了介绍。

Pituch et al（2009）、Preacher（2011）分别给出了基于MLM和MSEM的三水平中介效应分析的模型框架。Li＆Beretvas（2013）、McNeish（2017）探讨了多水平中介效应分析的样本量问题。Tofigh et al（2013）、Tofighi＆Kelley（2016）、Talloen（2016）评估了多水平中介效应分析中省略混淆因素偏误（omitted confounder bias）的问题。Preacher个人网站，含有关中介效应与调节效应分析的丰富资源，可了解该领域的前沿信息。有关多水平调节效应模型进展，参见Preacher et al（2016）。

多水平中介效应分析的实例如Hwang＆Hopkins（2012）、Nohe et al（2013）、Tofighi＆Thoemmes（2014）、Akremi et al（2014）、Oshio（2016）等。

---

(1) 当然，在所谓“上行效应”（bottom-up effects）（Bliese，2000；Kozlowski＆Klein，2000）或“微-宏观／突生效应”（micro-macro or emergent effects）（Croon＆van Veldhoven，2007；Snijders＆Bosker，1999）的研究中，这类模型也有现实性。实例可参见Nohe et al（2013）的研

究，使用的是2-1-1-2模型。Preacher et al（2010）提到了基于MSEM的1-1-2中介效应模型实例。本章也提供了2-1-2、1-2-1、1-1-2、1-2-2模型示例代码，见本章第4节 。

(2) 上标数字代表步骤，下同。这里实际上可以看出来，基于MLM进行多水平中介效应分析，本质上还是可以理解为区分步骤进行，可看作分别构建和估计了有关*M* 和引入*M* 后有关Y的多水平模型（因为一般基于ab乘积项进行估计，第一步用不上），只不过在M*plus* 中可以联立估计（相当于一种特殊SEM）。而基于MSEM法，则是结构性的估计方法，包括直接效应和间接效应各系数在作为整体的结构方程中同时估计（Kline，2011）。由此角度来讲，基于MLM和MSEM的多水平中介效应分析，可类比于单水平中介效应分别使用所谓层次回归（hierarchical regression）和路径分析模型进行估计。

(3) R软件中RMediation包可提供两个或多个系数乘积项区间估计及制图（Tofighi＆Mackinnon，2011）。

(4) 对于此类型有中介的调节效应（见图9-11），Van Kollenburg＆Croon（2017）将W称为proposed，initial moderator，M称为true，mediating moderator；W→（X→Y）、M→（X→Y）称为moderating effects；W→M→Y路径称为mediated moderating；X→Y称为moderated effect。

(5) 注意按照刘东等（2012）的方式，被中介的调节效应定义为ind＝$a×bb$ ，实际上这是“纯粹”意义的有中介的调节效应，也就是交互项系数乘上第二阶段的b路径系数（存在于组间）。读者可以参照一般单水平有中介的调节效应（参见王孟成，2014：48-50）进行重新界定，将变量的主效应也包括进来。

第三部分　贝叶斯结构方程模型

10　潜变量建模的贝叶斯方法

社会科学及行为科学领域，最常用的统计分析方法属于频率论或频率学派（例如，虚无假设检验）。在统计学传统里还有一个与之平行的统计学流派或思想称为贝叶斯学派。两种统计思想有着不同的理论渊源和哲学信念。在目前的应用统计实践中，频率学派占绝对的主导地位，但其思想的缺陷和不足也受到学界广泛批评（Cohen，1994）。近年来，社会科学各领域的学者开始呼吁重视贝叶斯统计（Depaoli＆van de Schoot，in press；Dienes，2011；Kruschke，Aguinis＆Joo，2012；Jackman，2004）。

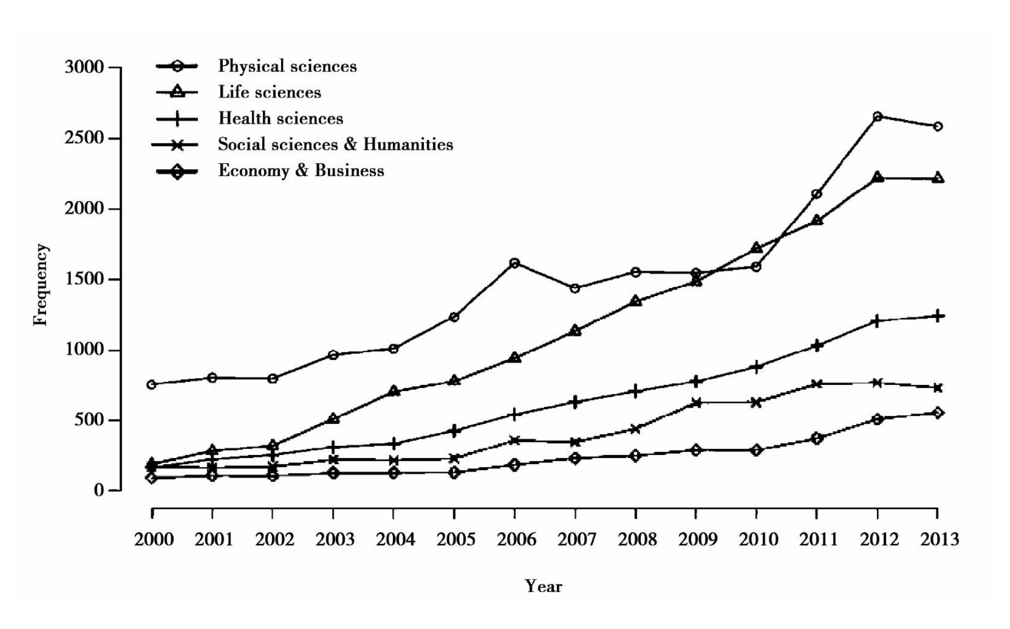


图10-1　贝叶斯统计在各学科近十几年的增长趋势（van de Schoot et al，in press）

图10-1是各大学科2000年以来使用贝叶斯方法所发表的论文数量及趋势。在应用领域，尽管贝叶斯统计并未撼动频率学派的统治地位，但其在各学科内的应用呈逐年稳步上升的态势。

潜变量建模使用贝叶斯方法有着许多优点，在有些情况下传统方法无法解决的问题对于贝叶斯来说轻而易举。本章主要分成两大部

分：第一部分主要介绍贝叶斯基本原理和分析过程。更细致详实的介绍可以参考Lynch（2010）和Gelman et al（2014）的专著。第二部分主要介绍几种常见的潜变量模型使用贝叶斯分析的过程和M*plus* 应用。

10.1　贝叶斯概述

贝叶斯与传统频率论之间最显著的区别在于对未知参数的定义。传统的频率统计将总体参数看成一个未知的固定值。比如，总体均值是一个固定不变的未知数。频率范式下的参数估计就是根据样本统计量去估计总体参数并建立相应的置信区间。与此不同，贝叶斯范式将参数看成一个不确定的变量，通常用一个分布来表达变量的可能取值范围。贝叶斯分析的目标就是确定这个分布（即后验分布）。通常使用均值和方差或标准差来描述一个分布的集中和离中趋势。同样地，后验分布的均值表示参数最可能的取值，类似点估计。后验分布的标准差类似于频率统计里经常使用的标准误。

贝叶斯将概率看作不确定性的主观信念。比如在掷骰子前人们总是根据自己的经验甚至“一厢情愿”的想法对结果做预测。生活中这样的例子比比皆是。比如，此时天空乌云密布、电闪雷鸣，根据以往的生活经验，我们判断接下来有95％～100％的概率会下暴雨。

在传统的频率统计里，我们常常会根据点估计和标准误构建总体参数的置信区间（Confidence Interval），即点估计±1.96／2.58标准误。贝叶斯统计也有置信区间，为了与传统的置信区间作区分，使用可信区间（Credible Interval）。尽管两个区间的形式类似，然而解释上存在极大的差异。具体来说，95％的置信区间，频率论的解释

是：如果重复从总体中抽取100个样本，其中有95个样本包含总体真值。而贝叶斯的可信区间是对概率的直接解释：总体真值有95%的可能性落在区间内。尽管我们经常将频率论的置信区间解释为贝叶斯的可信区间，但通过对比还是很容易看出两者的区别。

表10-1对比了贝叶斯和频率论之间的差异。

表10-1　贝叶斯与频率论差异比较

	贝叶斯	频率论
P 值	关于参数的概率	
认识论	承认过往研究的结果并以此为基础结合新样本更新知识	只从特定的样本中获取知识
样本要求	各种样本量	通常为大样本
参数的性质	未知随机值	固定值
总体参数的特点	不确定的分布	具体的值
估计区间的意义	区间包含 95%的总体值	从总体中抽取无数个样本，其中95%的样本包含总体真值

10.2　贝叶斯定理

假设两个独立事件A 和B 的概率记作$P(A)$ 和$P(B)$ 。根据概率理论，A、B 同时发生的概率等于条件概率乘以边际概率：

$$P(A, B)=P(A|B)\ P(B)=P(B|A)\ P(A) \qquad (10.1)$$

两边同除以$P(A)$ 后公式变成：

$$P(B|A)=\frac{P(A|B)P(B)}{P(A)} \qquad (10.2)$$

将上式稍加变化可以得到贝叶斯后验概率公式：

$$P(\theta \mid Y)=\frac{P(Y \mid \theta)P(\theta)}{P(Y)} \tag{10.3}$$

$P(\theta / Y)$指未知参数θ的后验分布，即在已知数据Y的条件下参数θ的条件分布。$P(Y \mid \theta)$是似然值，在假设参数的条件下数据的条件概率；$P(\theta)$为参数的先验分布；$P(Y)$为样本数据的边际分布。换句话说，后验概率等于先验概率乘以数据的似然值，因此上式也可以改写成：

$$\text{posterior}=\frac{\text{priors}\times\text{likelihood}}{\text{data}} \tag{10.4}$$

$P(Y)$分布的积分等于1，当Y分别是连续变量和分类变量时：

$$P(Y)=\sum_{\theta} P(Y \mid \theta)P(\theta) \tag{10.5}$$

$$P(Y)=\int_{\theta} P(Y \mid \theta)P(\theta) \tag{10.6}$$

通常$P(Y)$并不涉及模型参数，因此常将其省去：

$$P(\theta \mid Y) \propto P(Y \mid \theta)P(\theta) \tag{10.7}$$

$\propto$读作“正比于”（proportional to）。

10.3 贝叶斯分析

贝叶斯分析包含3个重要的步骤：（1）根据先前研究的结果或专家意见为未知参数设置先验分布；（2）通过似然函数计算样本数据的

结果；（3）根据贝叶斯定理将先验分布和似然函数结合形成参数的后验分布。

先验分布 [(1)]

先验信息（Prior）是指在数据收集前就已经获得的关于参数的信息、知识或信念。这些信息可能来自过往研究或元分析，也可能是研究者的经验，甚至是研究者的主观猜测。因此，根据先验分布富含信息的程度可以将其分为3种类型：有信息先验（Informative Prior）、无信息先验（Noninformative Prior or Diffuse Prior）和介于两者之间的部分信息先验（Weak informative Prior）。

先验分布方差的大小体现了研究者对先验信息的确定程度，如果对参数的可能结果非常有把握，可以设置一个小的方差。相反，如果对参数的可能取值范围并不是特别有把握，可以选择一个较大的方差以避免犯错。上述3种类别的先验信息就是通过设定不同的方差来区分的。

通常在研究之前已经存在关于某种现象或研究假设的知识，这些知识可以是前人的研究发现，也可以来自元分析（Meta-analysis）的结论。有信息先验可以帮助研究者在具体分析时借鉴前人的研究发现，特别是样本量较少的情况下，有信息先验可以弥补样本量少、抽样误差大的不足。如果先验信息是非常确定的，我们可以设定一个非常精确的取值范围。例如，某两个变量之间的相关，前人报告的系数范围为0.5～0.8，那么就可以将先验分布设定在这个区间，如N（0.6，0.1^2 ）。但很多时候，先验信息很难做到非常精确，由于错

误的设定先验信息可能会误导结果，所以会设定一个较大的范围，比如N（0.6，1^2 ）。

在缺少先验信息的情况下，可以设置无信息先验，通常设定非常大的方差[(2)]，例如10^{10} ，来体现不确定性。此时由于先验信息几乎没有提供任何信息，所以后验分布的结果主要取决于数据的信息，所得结果与极大似然估计的结果相近（Muthén＆Asparouhov，2012）。

3种不同的先验信息对后验分布的影响可以通过图形来表达。图10-2设置了3种不同的先验信息，分别是：（1a）无信息先验，方差非常大，为100^2 ，所以图中呈一条水平线。（2a）部分信息先验，方差比较大，为6^2 ，在图中表现为非常开阔。（3a）有信息的先验，方差比较小，为2^2 ，分布的形状相对高耸狭窄。数据的似然值是相同的，但是产生3个形状差异的后验分布：（1c）分布形态与数据似然值分布非常接近，说明无信息先验对后验分布产生微不足道的影响。（2c）与无信息先验相比，部分信息先验产生的后验分布其形态发生了显而易见的变化，坐标轴的位置也向左侧偏移，但整体来说后验分布和似然值更相似。（3c）有信息先验的后验分布的形态和位置更接近先验分布的形态和位置。

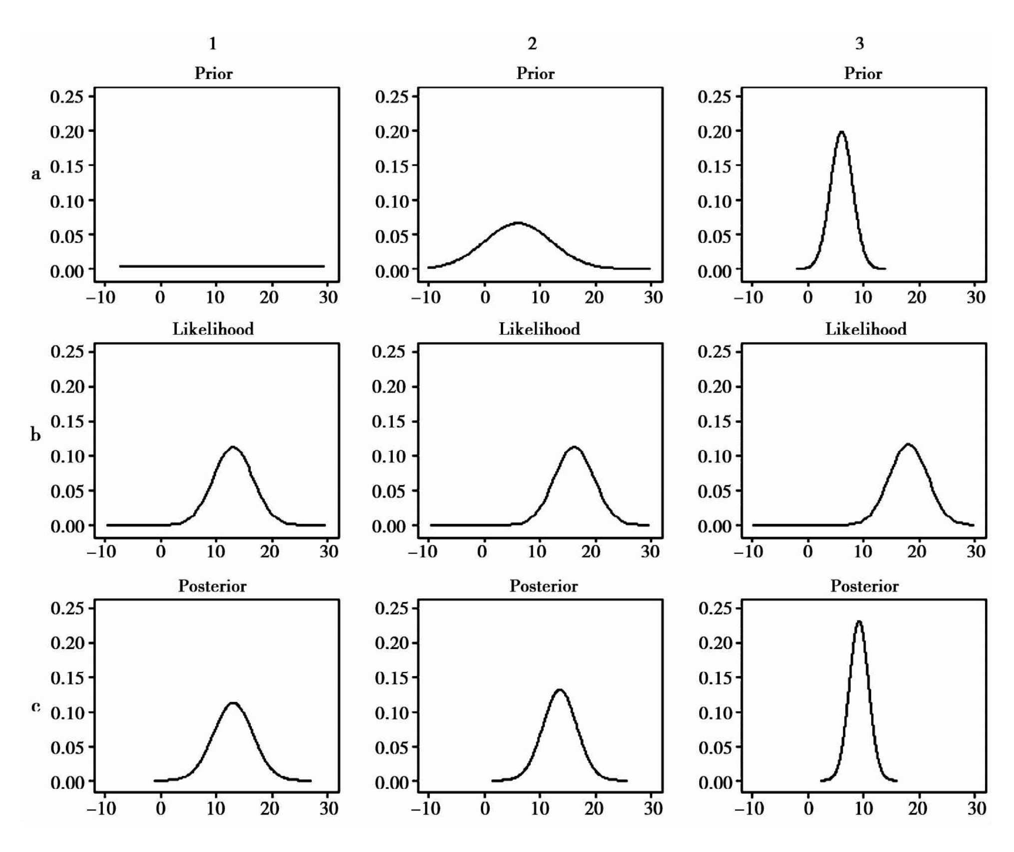


图10-2　不同先验信息对后验分布的影响示意图（Yuan＆MacKinnon，2009）

通过对同一数据设置3个不同的先验分布，不难发现所产生的3个后验分布被先验信息不同程度地“收缩”了：先验分布越“狭窄”（确定性越大）后验分布收缩越明显。

由于先验信息会对后验分布产生不同程度的影响，先验信息的设定是贝叶斯方法的关键和难点（MacCallum et al，2012；Yuan＆MacKinnon，2009）。如果设置的先验信息不合理会严重误导后验分布，导致完全错误的结论。正因为如此，很多方法学者对正确使用贝叶斯建模持谨慎态度（MacCallum，Edwards＆Cai，2012）。为了避免先验分布误设，可以使用敏感性分析（sensitivity analysis），即尝试采用不同的先验信息，观察其对后验分布的影响（e．g．，Gelman et al，2004）。

10.4 贝叶斯分析过程

根据以往研究结果确定先验分布，然后根据贝叶斯定理，将新数据加入先验分布得到参数的后验分布。这一过程是贝叶斯分析的过程，也是人类知识不断积累和更新的过程[(3)]：从前人的研究发现出发（先验知识或先验分布），设计新的研究收集数据（数据的信息即似然值），两者结合产生新的知识（后验分布）。而在传统的频率论中，即使前人对某一问题已经开展了很多相关研究并获得了一些发现（例如，X 正向预测 Y ，回归系数为0.2～0.3），但后来的研究仅仅根据前人的发现提出相同的假设，即 X 正向预测 Y ，但是之前的研究发现（系数为0.2～0.3）并没有被采纳。从这个意义上来说，频率论的统计思想阻碍了人类知识的积累。

贝叶斯学派作为统计学的两大学派之一，其在应用统计领域的影响越来越大，潜变量建模的很多情况下都可以使用贝叶斯方法来估计。在潜变量建模时，与传统方法相比贝叶斯估计有如下几个突出优点：

①小样本表现优良；

②计算速度更快；

③适应模型更多。

传统的估计方法不能识别或无法估计的模型，贝叶斯可以给出合理的结果。

10.4.1　贝叶斯估计：马尔科夫链蒙特卡洛算法

得益于马尔科夫链蒙特卡洛算法（Markov Chain Monte Carlo，MCMC），贝叶斯统计得以迅速发展和普及（Gelman et al，2004）。MCMC是一个迭代算法，首先设置先验分布，接着经过多次迭代产生参数的后验分布，迭代的过程形成一个马尔科夫链（chain）。MCMC至少需要两个开始点即两条马尔科夫链，以确保迭代过程收敛到一个稳定的后验估计。在实践中，为了确保收敛的稳定性，可以设置多个链并结合图形来判断。目前MCMC常采用吉布斯抽样（Gibbs sampler），其过程如下：

假设 θ_t 表示第 t 次迭代的全部模型参数（例如，因子分析里的因子负荷，因子协方差，误差方差等参数），吉布斯抽样将 θ_t 分成 d 组：$\theta_t = (\theta_{1t}, \cdots, \theta_{dt})$

第1步（$\theta_{1t} | \theta_{2t-1}, \cdots, \theta_{dt-1}$，*data, priors*）

第2步（$\theta_{2t} | \theta_{1t}, \theta_{3t-1}, \cdots, \theta_{dt-1}$，*data, priors*）

⋮

第 d 步（$\theta_{dt} | \theta_{1t}, \cdots, \theta_{d-1,t-1}$，*data, priors*）

循环上述步骤。

下面用文字简单描述这一过程。假设模型参数 θ 有5个（组）参数 θ_1 θ_2 θ_3 θ_4 θ_5，设定参数 θ 的初始值后，在给定 θ_2

θ_3 θ_4 θ_5 的条件下抽取 θ_1 的条件分布，产生第一个样本；接着在给定 θ_1 θ_3 θ_4 θ_5 的条件下抽取 θ_2 的条件分布样本，依次抽取5个样本。当所有的5个参数都更新（update）后，重复上述步骤，反复N 次迭代后，所有参数不再变化即可判断参数收敛了。

使用吉布斯抽样有两个重要问题：确定马尔科夫链数和迭代次数。至于马尔科夫链的个数通常采用多个，M*plus* 默认为两个。设置多个链一来可以提高计算速度（使用多核处理器），另外也可以用于判断收敛，比如下面提到的PSR。当多个链的结果都收敛到一个稳定的分布时（模型收敛了），后验统计量就可以计算了。但是每个链收敛的过程可能差异较大，而且在迭代的前期链内部的变异较大，如果基于整个迭代过程计算后验统计量会使结果产生偏差。因此，通常会将迭代过程分成前后两部分，前半部分的结果丢掉不用，即所谓的“煲机”或热身阶段（burn-in），只使用后半部分的迭代结果计算后验统计量。

贝叶斯估计也会存在类似极大似然估计收敛到局部解的问题。在采用极大似然估计分析混合模型时，通常使用多个起始值来避免局部最大化解，设定多个马尔科夫链与设定多个起始值类似。然而马尔科夫链的长度是由迭代次数决定的，所以确定迭代次数是判断收敛的关键。如果迭代次数太少，结果可能收敛到一个“局部”的后验分布，如图10-3下方所示。上方图是迭代6000次的马尔科夫链，结果很理想，但当迭代次数增加至50000次时，结果发生逆转，而真正的收敛区域在6000次迭代后。

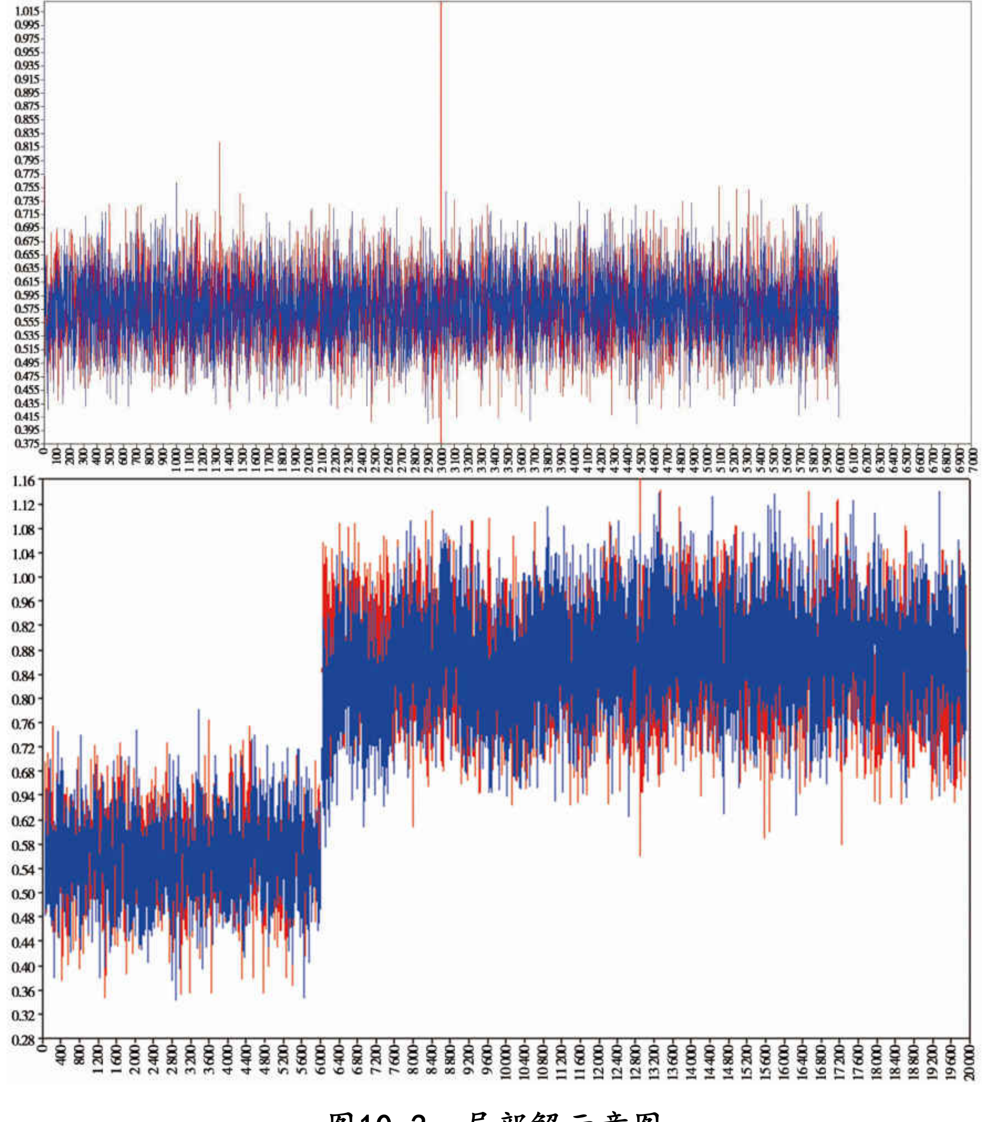


图10-3　局部解示意图

为了避免局部解，Depaoli等最近建议先使用一个固定迭代次数，例如50000次，如果结果收敛了，再使用加倍的迭代次数即100000次分析，如果两次的结果差异不大则说明没有出现局部解。由于局部解很难事先判断，所以最好的做法是尝试不断地加倍迭代次数，考虑到现在计算机的高性能，在时耗上这种策略也是可以接受的[4]。

10.4.2　收敛的判断

收敛是指马尔科夫链是否达到平稳状态，如果达到了平稳状态则说明得到的样本是从目标抽样分布中抽取的样本。

由于MCMC估计时参数收敛的是一个分布而非一个固定的点，所以判断参数收敛是很困难的。由于不存在单一的统计量用以评价此种情况，贝叶斯模型中判断收敛的标准有多个，其中最常用的是对MCMC链收敛过程的图形检验即踪迹图（trance plot）。另外，常用的还有自相关图（autocorrelation plot）、PSR和K-S检验，下面分别介绍这些方法。

（1）踪迹图

如果将多个马尔科夫链同时呈现在一张图中（软件通常使用不同的颜色标记不同的链），就可以直观地判断不同的马尔科夫链是否收敛到一个相同的平稳分布，如图10-4所示。横坐标是迭代次数，纵坐标是后验估计值，在图的中心位置出现一条纵线将踪迹图分为前后相等两部分，前半部分是“burn-in”阶段，通常波动较大，后半部分结果相对稳定，用作模型收敛的判断依据。

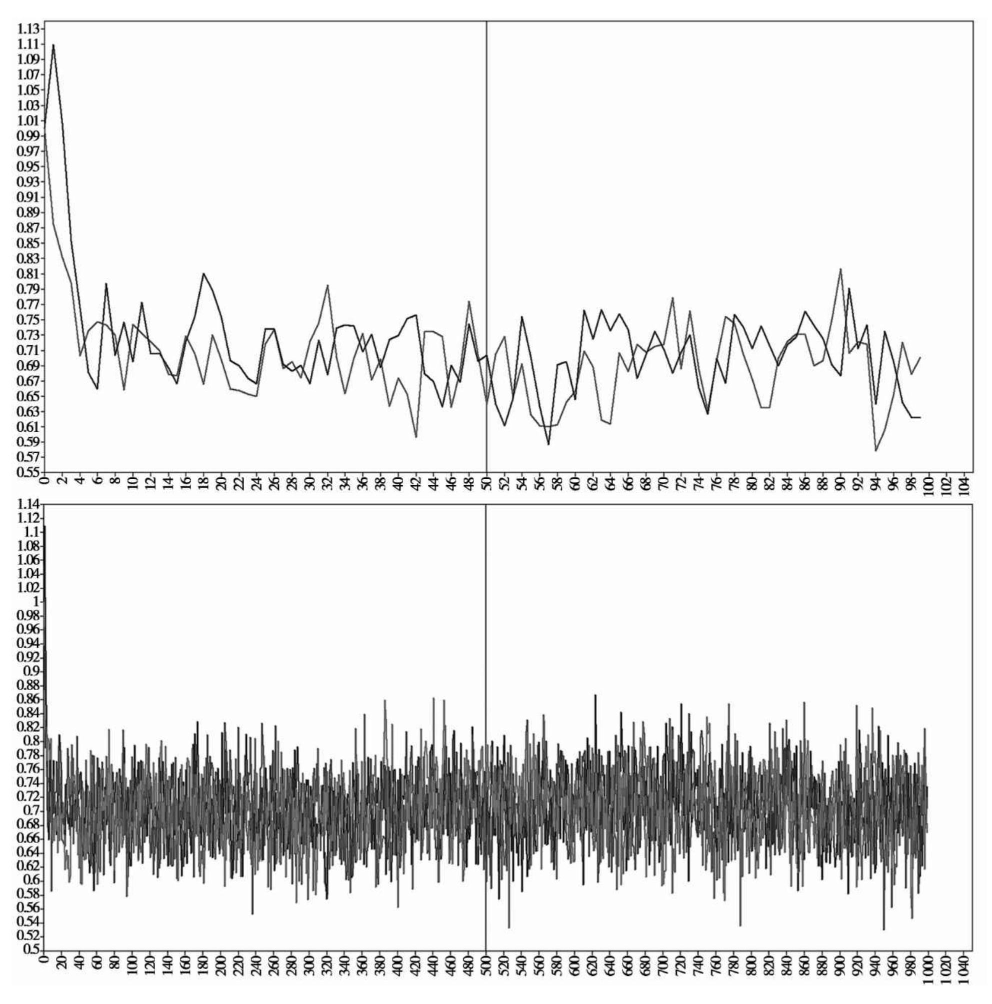


图10-4　踪迹图（左图有2条链100次迭代；右图有4条链10000次迭代）

（2）自相关图

根据MCMC抽样的特性，每次抽样只依赖于前次抽样，因此可以计算每次抽样参数间的相关系数。理论上非连续抽样之间不相关，但实际中很难实现，所以实践中通常计算间隔一定距离的抽样相关，比如每间隔10次抽样，这个间隔称为筛选间隔或抽样步长（thinning interval or sampling lag）。如果自相关系数小于0.1，说明抽样样本接近独立。图10-5是一个自相关系数分布图，纵坐标是相关系数大

小，横坐标是不同的迭代间隔次数或步长。如果自相关系数较大而且随着步长的增加相关系数会逐步减少，可以使用M*plus* 的Thin语句只报告间隔一定步长的结果（Muthén，2010）。

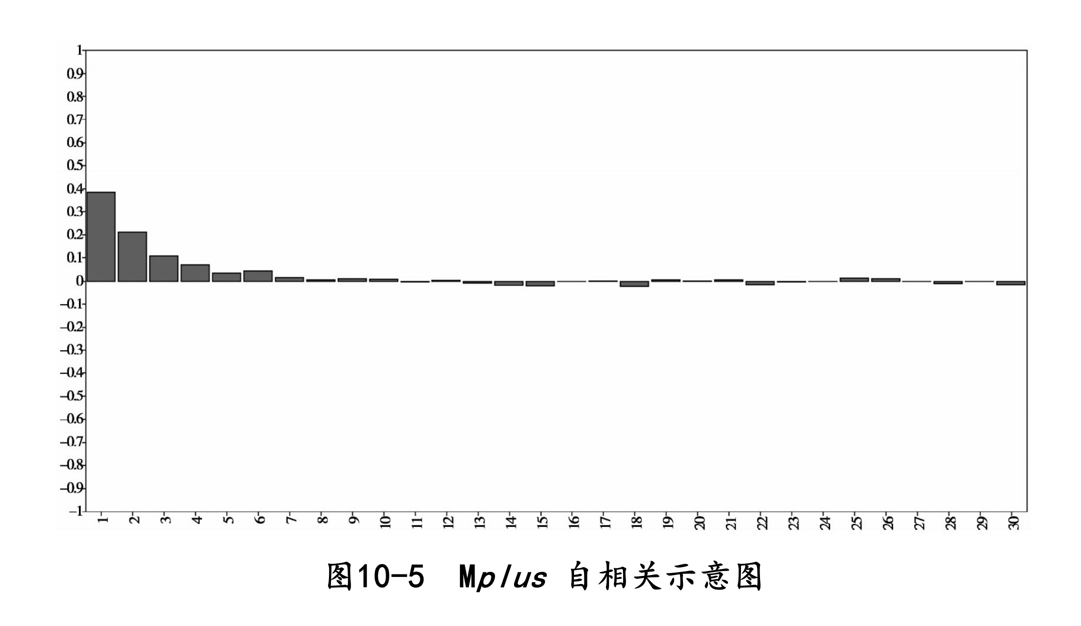


图10-5　M*plus* 自相关示意图

（3）潜在尺度缩减因子

收敛可以通过潜在尺度缩减因子（Potential Scale Reduction，PSR；Gelman et al，2014，第11章），也称为Gelman-Rubin法，来判断。

令 θ 表示模型中的参数（例如，方差、均值等），$\bar{\theta}_j$ 表示第j 条链的均值；θ_{ij} 表示第j 条链上的第i 次迭代的值；$\bar{\theta}$ 表示所有链的总均值，PSR计算公式如下：

$$\overline{\theta_j} = \frac{1}{n}\sum_{i=1}^{n}\theta_{ij} \tag{10.8}$$

$$\bar{\theta} = \frac{1}{m}\sum_{j=1}^{m}\overline{\theta_j} \tag{10.9}$$

$$B = \frac{1}{m-1}\sum_{j=1}^{m}(\bar{\theta}_j - \bar{\theta})^2 \tag{10.10}$$

$$W = \frac{1}{m}\sum_{j=1}^{m}\frac{1}{n}\sum_{i=1}^{n}(\bar{\theta}_{ij} - \bar{\theta}_j)^2 \tag{10.11}$$

$$\mathrm{PSR} = \sqrt{\frac{W+B}{w}} \tag{10.12}$$

PSR的核心思想是：当不同链之间的变异小于链内变异时，说明不同链间的结果趋于相等，模型收敛了[5]。PSR也可以通过将不同链的踪迹图画在同一个图中来判断。

PSR类似组内相关系数ICC，但又有所不同。它与ICC存在如下关系：

$$\mathrm{ICC} = \frac{B}{W+B} = 1 - \frac{1}{\mathrm{PSR}^2} \tag{10.13}$$

（4）K-S检验

K-S检验是用于比较连续的单维概率分布是否相等的非参数检验方法，比PSR更严格，在M*plus* 中可以在TECH8中获得每个参数KS检验结果。如果检验的*p* 值非常小，软件会给出如下提示：

THE KOLMOGOROV-SMIRNOV DISTRIBUTION TEST FOR PARAMETER 52 HAS A P-VALUE 0.0000，INDICATING DISCREPANT POSTERIOR DISTRIBUTIONS IN THE DIFFERENT MCMC CHAINS.

THIS MAY INDICATE NON-CONVERGENCE DUE TO AN INSUFFICIENT NUMBER OF MCMC ITERATIONS OR IT MAY INDICATE A NON-IDENTIFIED MODEL. SPECIFY A LARGER NUMBER OF MCMC ITERATIONS USING THE FBITER OPTION OF THE ANALYSIS COMMAND TO INVESTIGATE THE PROBLEM.

10.5　模型拟合评价与模型比较

10.5.1　后验预测检验

后验预测检验（Posterior Predictive Checking，PPC；Gelman et al，1996）是模型评价最常采用的方法。后验预测检验背后的思想是：如果模型拟合数据，那么基于该模型产生的随机样本应该和观测数据差不多。换句话说，比较从联合后验预测分布中抽取的随机样本和观测数据，如果发现两者存在系统差异则说明模型拟合出现问题。

后验预测检验的p 值是通过假设模型与自由估计模型的卡方似然比检验获得的（p 值表示观测数据卡方值比模型生成数据卡方值小的比例）。p 值越小模型拟合越差，例如p ＝0.05（表示观测数据卡方值比模型生成数据卡方值小的比例只有5％）。当模型拟合良好时后验预测p 值应该接近0.5，表明观测数据和根据模型生成的数据之间差不多。需要注意的是：

①这个p 值和传统显著性检验的p 值不同，并不意味着正确模型的一类错误率为5%，后验预测p 值更像拟合指数而非显著性水平。

②该值只能用来评价单个模型的拟合情况，不能用来进行模型比较（Meng，1994）。

在M*plus* 中默认的后验预测p 值通过计算所有第10次迭代获得（Asparouhov＆Muthén，2010）。另外，M*plus* 还可以给出PPC的分布直方图（图10-6）和散点图（图10-7）。

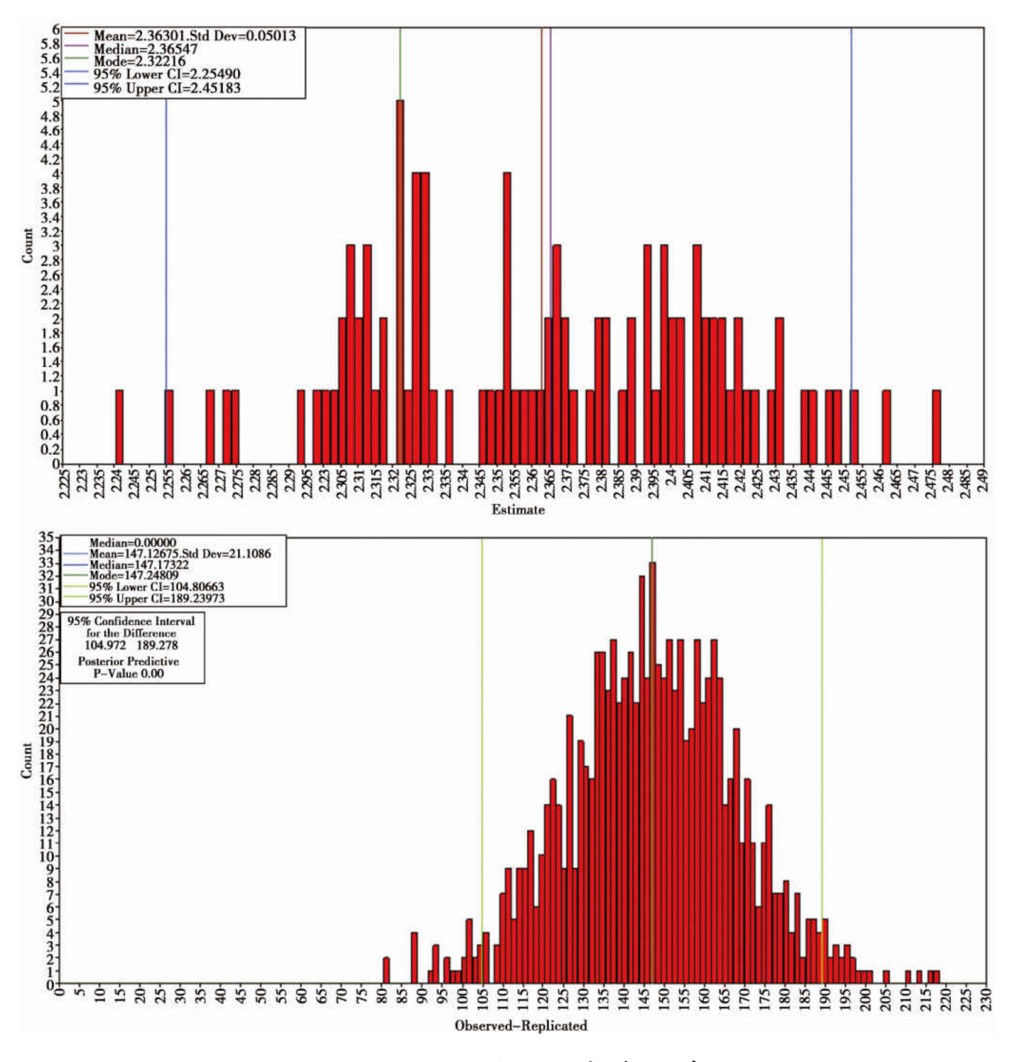


图10-6　后验预测检验分布的直方图

图10-6中的竖线是观测数据结果，柱状图形成的分布是根据模型抽取样本的统计量形成的分布（纵坐标是结果计数，横坐标是估计结果）。如果观测数据结果（竖线）在分布图的中间位置，说明模型拟合数据较好。图10-7中的散点图与之类似，如果模型拟合数据良好，散点会落在45°线上，而图中的散点绝大多数落在斜线上，说明模型拟合良好。直方图和散点图的结果与PPC的p 值是一致的。

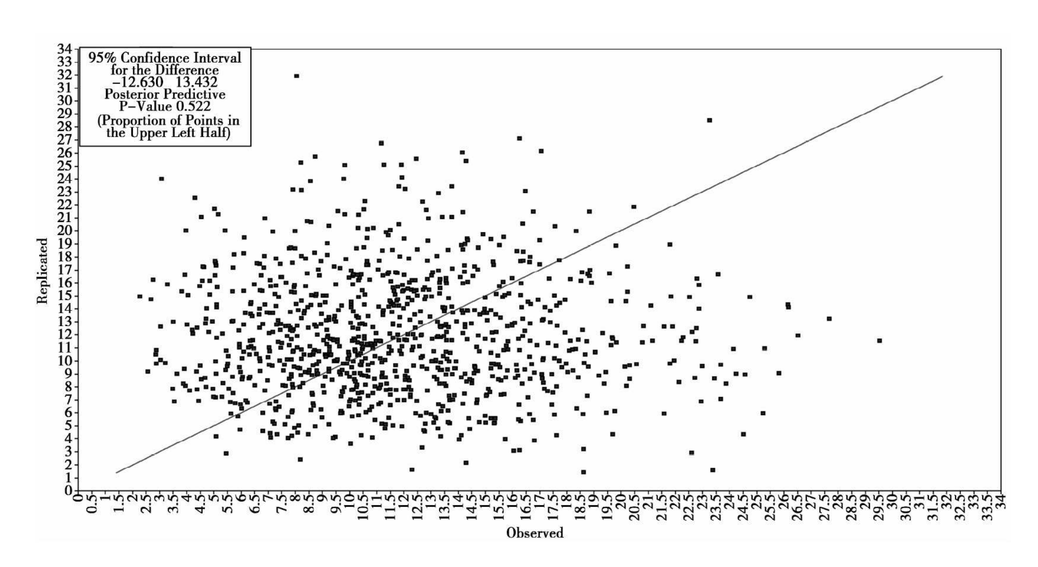


图10-7　后验预测检验分布的散点图

如果参数的后验分布直方图呈现不连续（如图10-6上方图所示）则提示迭代次数不足（Depaoli＆van de Schoot，in press），需要增加迭代次数。本例中将迭代次数增加到10000次后，结果如图10-6的下方图所示。

10.5.2　贝叶斯因子

贝叶斯因子（Bayes Factors；Kass＆Raftery，1995）是模型比较常用的指标。该指标即适用于嵌套模型也适用于非嵌套模型比较。

假设存在两个竞争模型M_1和M_2，其参数向量分别记作θ_1和θ_2，根据贝叶斯定理，M_1的后验概率：

$$p(M_1|y)=\frac{p(y|M_1)p(M_1)}{p(y|M_1)p(M_1)+p(y|M_2)p(M_2)} \tag{10.14}$$

其中下式表示已知M_1时y的边际概率：

$$p(y|M_1)=\int p(y|\theta_1,M_1)p(y|\theta_1,M_1)d\theta_1 \tag{10.15}$$

M_1和M_2的后验概率odd比可以写成：

$$\frac{p(M_1|y)}{p(M_2|y)}=\left[\frac{p(y|M_1)}{p(y|M_2)}\right]\times\left[\frac{p(M_1)}{p(M_2)}\right] \tag{10.16}$$

其中前半部分即为贝叶斯因子：

$$BF=\frac{p(y|M_1)}{p(y|M_2)}=\frac{\int p(y|\theta_1,M_1)p(\theta_1|M_1)d\theta_1}{\int p(y|\theta_2,M_2)p(\theta_2|M_2)d\theta_2} \tag{10.17}$$

Kass＆Raftery（1995）给出了模型比较时贝叶斯因子的经验法则，见表10-2。

表10-2　贝叶斯因子临界值

BF	BIC差值	拒绝H_2或M_2
1～3	0～2	不确定

3～20	2～6	支持
20～150	6～10	强
＞150	＞10	非常强的支持

当两个竞争模型的贝叶斯因子小于3时，两个模型差不多；当贝叶斯因子为3～20时，数据结果提示支持M_1；当贝叶斯因子为20～150时，说明数据结果“强力”支持M_1；当贝叶斯因子大于150时，数据结果“非常确定”地支持M_1。

10.5.3　贝叶斯信息指数

BIC（Bayesian information criterion，BIC）是频率统计和贝叶斯方法通用的模型比较的指标，最早由Raftery（1986）提出。

$$BIC = -2\log L + t\log N \tag{10.18}$$

式中N为样本量，t为自由参数的个数，$\log L$为对数似然值。$-2\log L$用于模型拟合，$t\log N$用于惩罚模型复杂性。由于考虑了样本量的影响，所以BIC倾向于支持更简单模型。

BIC可用于嵌套和非嵌套模型的比较，通常采用两个模型的差值来衡量：$\Delta BIC = BIC_1 - BIC_2$。$\Delta$BIC与贝叶斯因子存在一定的对应关系，见表10-2。

10.5.4　异常信息指数

异常信息指数（Deviance Information Criterion，DIC；Spiegelhalter，Best，Carlin＆van der Linde，2002）是评价模型拟合和竞争模型比较常用的指标。DIC可以看作AIC的一般化。

$$D_{(\theta)} = -2\log[p(y \mid \theta)] + 2\log[h(y)] \tag{10.19}$$

$$\overline{D_{(\theta)}} = E_{\theta}[-2\log[p(y \mid \theta) \mid y] + 2\log[h(y)]] \tag{10.20}$$

式中的h（y）是标准化因子，不参与模型比较（Spiegelhalter et al，2002）。

与BIC类似，DIC小的模型拟合更好。进行模型比较时，如果DIC差异小于5，而且两个竞争模型差异较大，选择DIC小的模型可能会得到错误的结论。

10.6　贝叶斯结果报告：WAMBS清单

WAMBS（*When to worry and how to Avoid the Misuse ofBayesian Statistics*；Depaoli＆van de Schoot，in press）清单是研究者最近提出的，旨在提高贝叶斯分析实践中结果的透明度和可重复性。这个清单包含4个方面的10个要点，具体内容罗列在表10-3中。

表10-3　WAMBS检核表（《何时需注意及怎样避免误用贝叶斯统计》）

	应该担心？	应该咨询专家？

模型估计前的核查要点		
第1点：是否理解先验？	是 / 否	是 / 否
估计后结果检查前的核查要点		
第2点：踪迹图是否收敛？	是 / 否	是 / 否
第3点：加倍迭代后是否收敛？	是 / 否	是 / 否
第4点：直方图是否有足够的信息？	是 / 否	不存在
第5点：自相关系数是否太大？	是 / 否	是 / 否
第6点：后验分布是否有实质性意义？	是 / 否	是 / 否
理解先验信息的影响		
第7点：设置不同的多元方差先验是否会影响结果？	是 / 否	是 / 否
第8点：与无信息先验相比，有信息先验是否显著影响结果？	从不	不存在
第9点：敏感性分析的结果是否依然稳定？	从不	是 / 否
模型结果解释之后		
第10点：是否使用了贝叶斯的结果解释和报告方式？同时也报告了缺失值、模型拟合和比较、概括化等。	是 / 否	是 / 否

该清单主要涉及以下4个方面的内容：

第一个方面关于研究者是否理解先验分布。

第二个方面涉及模型收敛评价。

第三个方面涉及设置的先验是否对后验分布产生影响即敏感性分析。

第四个方面是结果报告的规范。

目前关于贝叶斯统计建模需要报告哪些内容并没有公认的标准，该清单的提出将会规范今后的研究报告格式。这个清单包含分析前后的整个过程，正如提出者所言“完成这个清单是件非常耗时、枯燥的事，但为了……”。不管将来国内学术期刊是否接受或采用这个清单，至少这个清单可以帮助我们理解贝叶斯分析的过程以及哪些环节至关重要。

10.7 贝叶斯CFA

10.7.1 背景介绍与模型设定

这里我们使用“基础篇”第5章验证性因素分析时使用的数据。测量模型的设定选用King的四因子结构。这里我们设定3种不同的先验信息（为了简便这里只设定了负荷的先验信息，其他参数当然也可以设定）：

（1）无信息先验

分析时选择贝叶斯估计（ESTIMATOR＝BAYES），其他设定采用软件默认，语句见框10-1。如果只选择贝叶斯估计而不指定参数的先验分布，软件将采取默认设定即无信息先验：所有截距和负荷先验分布设定为正态分布均值为0，方差无穷大；误差方差设定为逆伽玛分布（inverse gamma，IG）；方差-协方差设定为逆威舍特（Inverse Wishart，IW）分布，这些默认的设定可以在TECH1中查到。M*plus* 默认的收敛依据是PSR，当所有参数的PSR小于1.05（单个参数）或1.1（多个参数）时迭代收敛。

框10-1　贝叶斯估计：无信息先验（M*plus* 默认估计）

```
TITLE:     The structure of PTSD of King's Four Factors using
Bayes
DATA:   File is PTSD. dat;
VARIABLE: NAMES ARE x1 x2 y1-y17;
    USEVARIABLES are y1-y17;
ANALYSIS: ESTIMATOR＝BAYES; ! 选择贝叶斯估计
    FBITER＝50000;
MODEL: F1 BY Y1-Y5*

; ! King's Four Factors model
    F2 BY Y6-Y7*

;
    F3 BY Y8-Y12*
```

```
;
    F4 BY Y13-Y17*


;
    F1-F4@1;
OUTPUT: STANDARDIZED TECH8;
PLOT:   TYPE＝PLOT2;
```

（2）信息先验

如果根据过往的研究可以获得条目负荷的大致范围，那么我们就可以设置有信息先验。假如过往的研究报告的因子负荷为0.4～0.8，平均为0.5，则此时我们可以将因子负荷的先验分布设定在一个相对确定的区间N（0.5，0.2），语句见框10-2。

框10-2　贝叶斯估计：有信息先验分布（自定义）

```
TITLE:     The structure of PTSD of King's Four Factors using
Bayes
DATA:   File is PTSD. dat;
VARIABLE:
NAMES ARE x1 x2 y1-y17;
          USEVARIABLES are y1-y17;
ANALYSIS:
```

```
ESTIMATOR＝BAYES；！选择贝叶斯估计
           FBITER＝50000；！固定迭代次数（将使软件不以PSR作为收敛标
准）
MODEL:    F1  BY Y1-Y5*

（a1-a5）；！King's Four Factors model
          F2 BY Y6-Y7*

（b1-b2）；
          F3 BY Y8-Y12*

（c1-c5）；
          F4 BY Y13-Y17*

（d1-d5）；
          F1-F4@1;
MODEL PRIORS：！设置模型先验信息（注意这里只设定了负荷的先验信息）
          A1-A5～N（0.5，.2）；！将0.2改成25就是部分信息先验的设置
          B1-B2～N（0.5，.2）；
          C1-C5～N（0.5，.2）；
          D1-D5～N（0.5，.2）；
OUTPUT:     STANDARDIZED TECH8;
PLOT:       TYPE＝PLOT2;
```

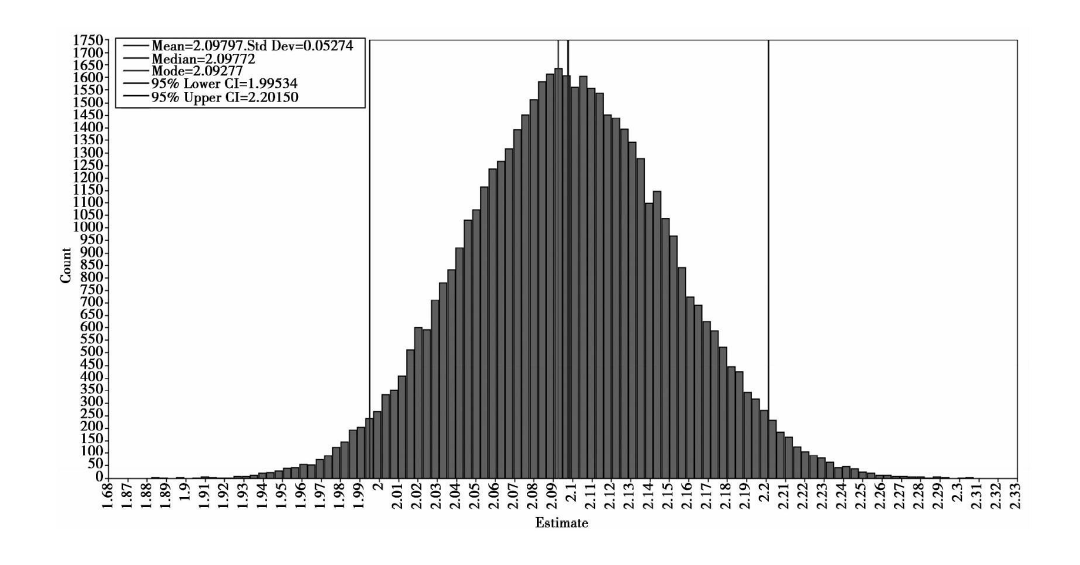
（3）部分信息先验

如果我们对负荷的范围并不是很确信，可以扩大方差为N（0.5，25）。

3种不同设定的负荷结果列在表10-4中。

10.7.2　模型参数收敛评价

模型迭代次数设定为50000次，前25000次为burn-in阶段。图10-8是条目16的负荷的后验预测分布、踪迹图和两条马尔科夫链对应的自相关图。



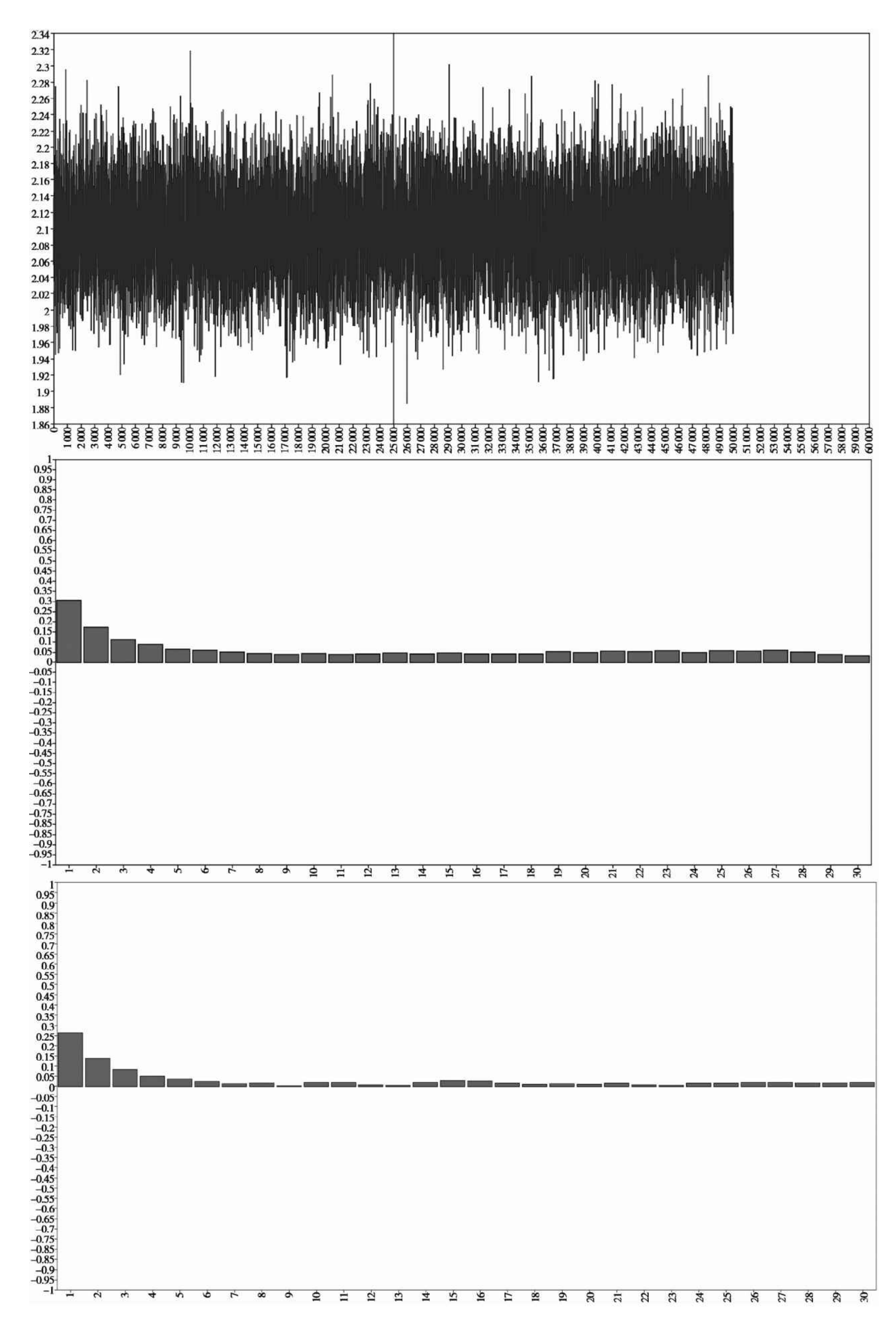


图10-8　条目16的后验预测分布、踪迹图和自相关图（两条链）

从踪迹图的结果来看，模型参数收敛到一个稳定的分布，自相关也在合理的区间。TECH8报告的PSR结果显示，第500次迭代时PSR＝

1.065，满足通常的收敛标准，33300次迭代之后结果稳定在1.001。这些结果提示，模型参数估计收敛良好。

10.7.3　模型拟合与比较

图10-9呈现了模型的后验预测直方图和散点图。图中表明了后验预测p 值等于0.000，说明模型拟合不理想，提示模型不能很好地复制观测数据。其中95％的可信区间为［103.964 190.762］，提示观测数据的卡方值大于后验分布生成数据的卡方值。

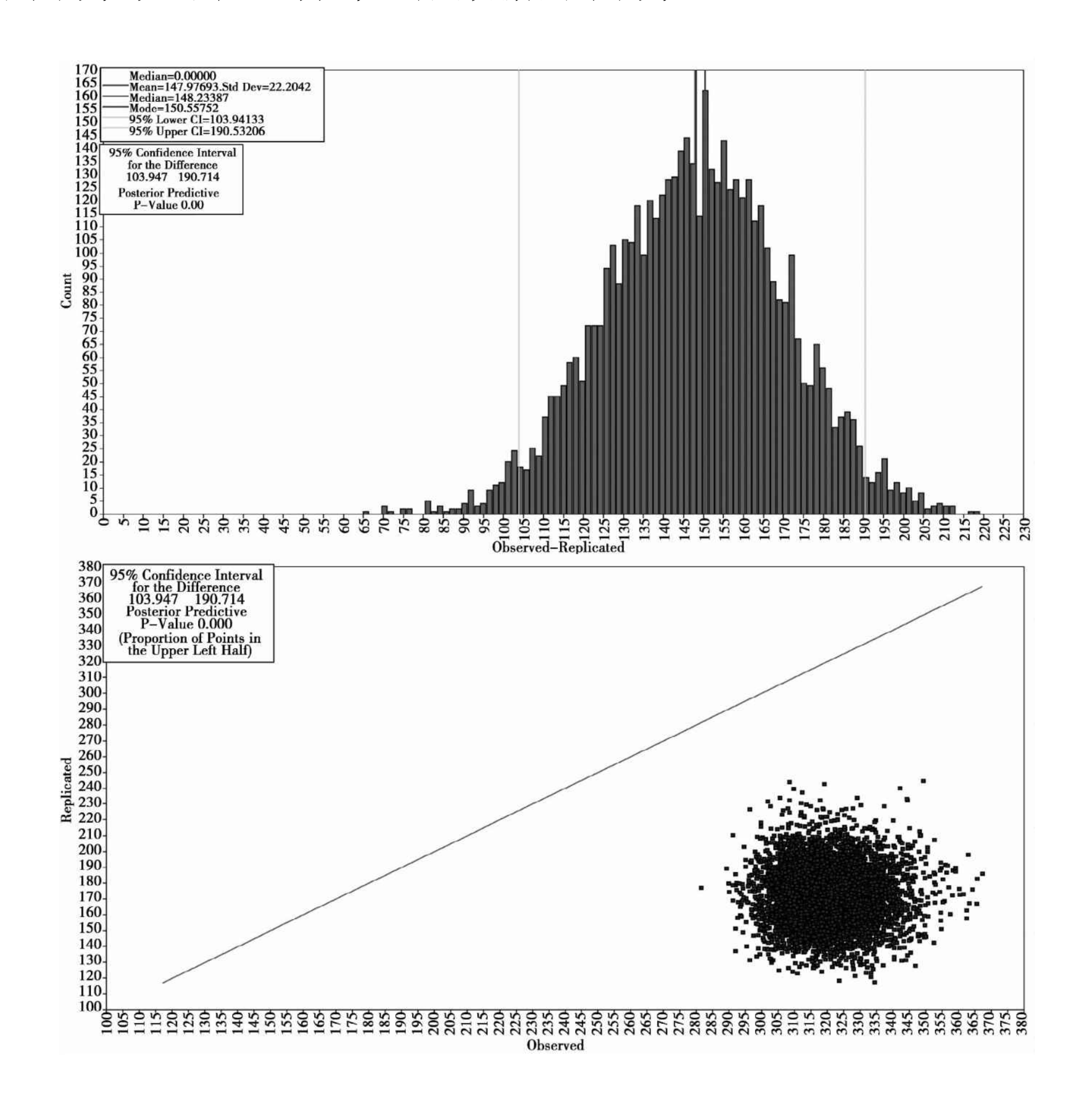


图10-9　CFA模型的后验PPC直方图和散点图

另外，King四因子模型的DIC＝27632.183，原DSM-4三因子模型对应的DIC值＝27750.595，两者相差118.412，提示四因子模型优于三因子模型。

10.7.4　模型结果解释

运行框10-1的语句后，软件输出的部分结果列在框10-3里。框中第一列数据为点估计结果，M*plus* 默认的是中位数估计。在有调节的中介效应模型的研究中发现均值估计比中位数估计更精确（Wang＆Preacher，2015），这一结果是否适合其他情形目前尚不明确，需要将来的研究对此进行验证。第二列是后验标准差。第三列数值是基于后验分布的单侧显著性检验*p* 值。如果参数估计结果是正值，*p* 值表示后验分布小于0的比例；如果参数估计是负值则表示后验分布大于0的比例（Muthén，2010）。最后两列数是95％的可信区间。

框10-3　M*plus* 输出结果示例

STDYX Standardization

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	Upper 2.5%	Significance
F1 BY						
Y1	0.571	0.034	0.000	0.502	0.633	*
Y2	0.491	0.036	0.000	0.417	0.558	*
Y3	0.711	0.028	0.000	0.653	0.761	*
Y4	0.601	0.032	0.000	0.534	0.659	*
Y5	0.632	0.030	0.000	0.571	0.688	*
F2 BY						
Y6	0.700	0.031	0.000	0.637	0.758	*
Y7	0.788	0.031	0.000	0.726	0.847	*
F3 BY						
Y8	0.597	0.032	0.000	0.531	0.656	*
Y9	0.637	0.030	0.000	0.574	0.693	*
Y10	0.757	0.026	0.000	0.703	0.804	*
Y11	0.671	0.028	0.000	0.613	0.723	*
Y12	0.472	0.037	0.000	0.397	0.542	*
F4 BY						
Y13	0.559	0.033	0.000	0.491	0.620	*
Y14	0.598	0.032	0.000	0.533	0.657	*
Y15	0.600	0.032	0.000	0.533	0.660	*
Y16	0.605	0.032	0.000	0.540	0.663	*
Y17	0.564	0.034	0.000	0.494	0.627	*

表10-4中是3种不同的先验分布所对应的后验结果（这里只有点估计的结果）。从表中结果不难发现，无信息先验的结果与传统的极大似然估计的结果非常接近，两者相差在0.001～0.003。两种有信息先验的估计结果与无信息先验的结果相差也不大，但可以发现这样一种趋势，即先验信息越明确，后验估计值的结果与极大似然估计的结果差异越大。

表10-4　不同先验分布的后验结果 [(6)]

Item	ML	Bayes(default)	Bayes N(0.5, 0.2)	Bayes N(0.5, 25)
A1	.571	.571	.557	.571
A2	.490	.491	.489	.491
A3	.711	.711	.691	.711
A4	.601	.601	.587	.601
A5	.632	.632	.627	.632
B1	.702	.700	.694	.700
B2	.791	.788	.780	.788
C1	.597	.597	.592	.597
C2	.637	.637	.630	.637
C3	.757	.757	.745	.757
C4	.670	.671	.668	.671
C5	.471	.472	.471	.472
D1	.559	.559	.549	.559
D2	.598	.598	.581	.598
D3	.600	.600	.577	.600
D4	.606	.605	.574	.605
D5	.565	.564	.542	.564

10.7.5　敏感性分析

根据WAMBS清单，需要检验不同的先验分布对后验估计的影响即敏感性分析（Depaoli＆van de Schoot，in press）。在这里我们设定了3种不同的先验信息（即无信息先验、部分信息先验和有信息先验），所得结果差异不大，因此可以得到敏感性分析结果依然稳定的结论。

10.8　贝叶斯中介模型

10.8.1　背景介绍和模型设定

中介分析是社会科学各领域应用最广泛的一种分析模型，我们在“基础篇”的第3章和第8章分别介绍了观测变量和潜变量的中介分析即路径分析和结构方程模型，使用的都是传统的频率论方法（即极大似然估计）。下面我们仍然采用这些例子来演示使用贝叶斯估计的分析过程，对比不同方法所得结果的差异。

这里的模型设定和传统的极大似然估计时的模型设定相比并无改变。

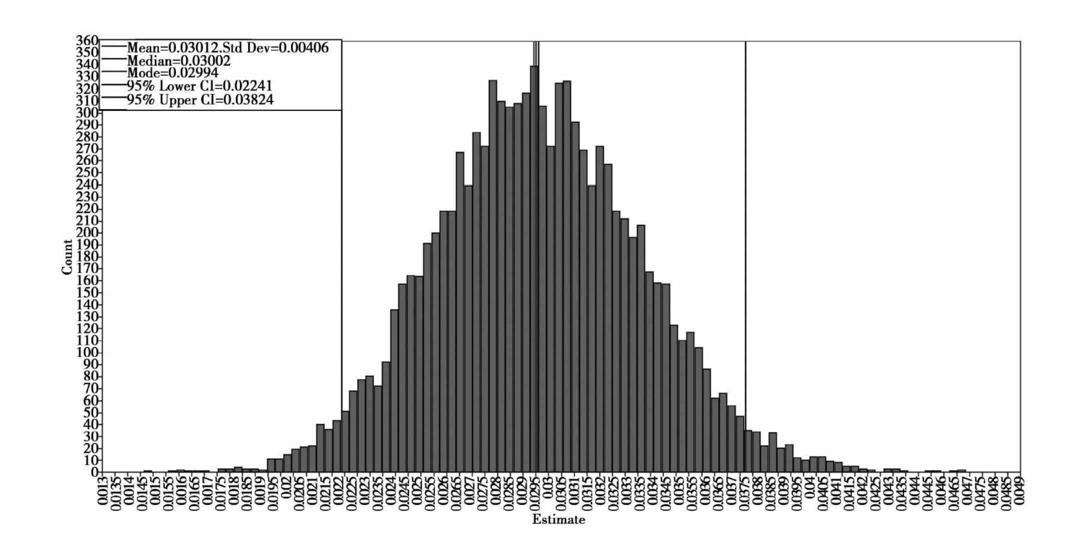
框10-4　贝叶斯中介分析

```
TITLE: this is an example of a path analysis
DATA: FILE IS 8-data. dat;
VARIABLE: NAMES ARE age gender a1-a5 e1-e13 b1-b20 c1-c17 d1-
d10;
  USEVARIABLE＝stress negtive depre self;
DEFINE: stress＝sum（a1-a5）;
    negtive＝sum（b13-b17）;
    depre＝sum（e1-e5）;
    self＝sum（d1-d5）;
ANALYSIS: ESTIMATOR＝BAYES;
FBITER＝10000;
chain＝2;
MODEL: depre on negtive stress self;
        negtive on stress;
        self on stress;
```

```
        negtive with self;
MODEL INDIRECT:
      depre ind self stress;
      depre ind negative stress;
OUTPUT:   standardized TECH8;
PLOT: TYPE＝PLOT2;
```

10.8.2　模型收敛检验

图10-10是中介模型估计结果对应的后验参数分布、踪迹图和自相关图（对应negative on stress路径系数）。这些结果提示模型参数收敛合理。



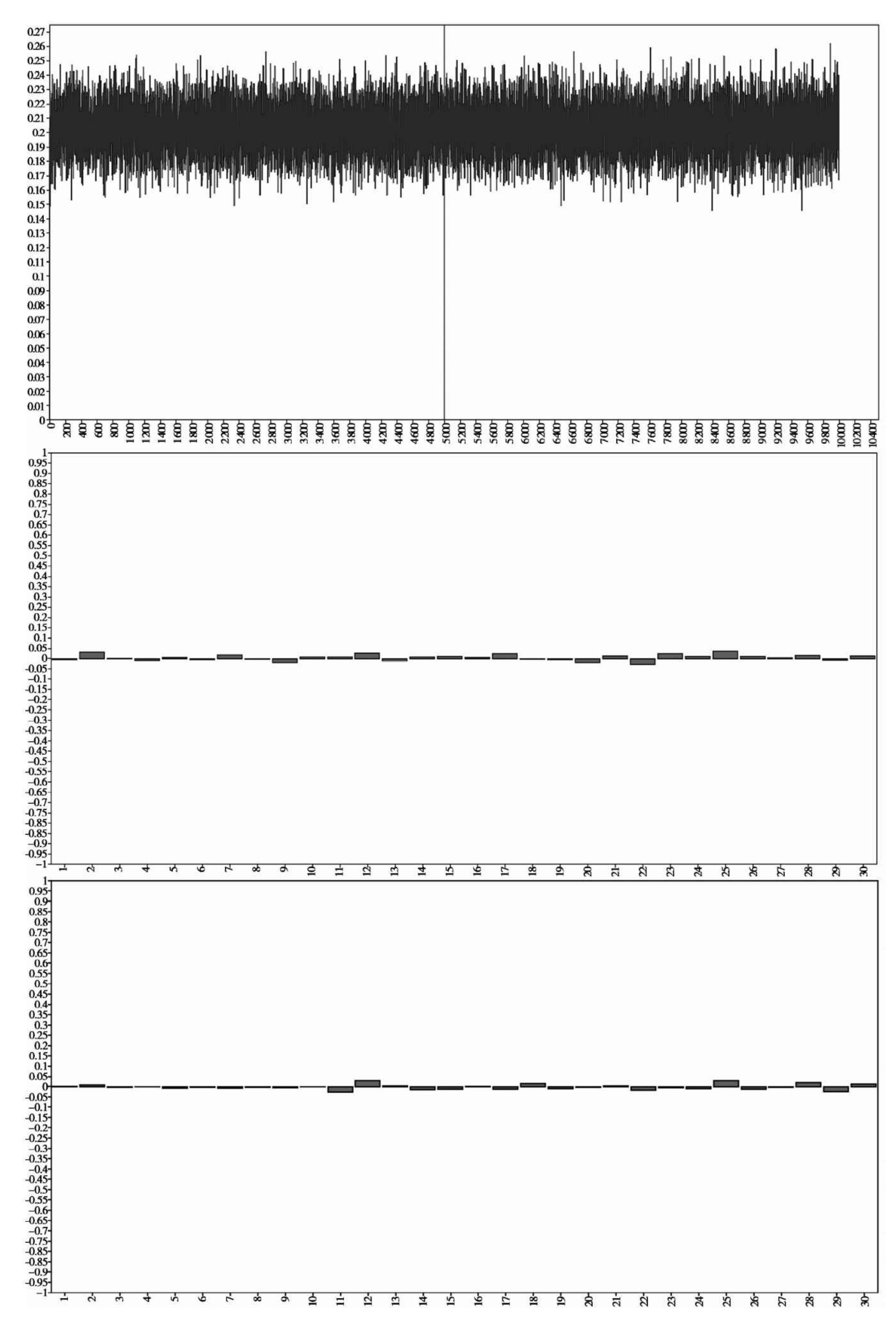


图10-10 贝叶斯中介分析的后验参数估计、踪迹图和自相关图 [7]

10.8.3 模型拟合评价

后验预测检验的p 值和标准化路径系数的输出结果呈现在框10-5中。其中PPP ＝0.522，对应的散点图和直方图见图10-11。按照前面介绍的评价标准可知模型拟合较好。

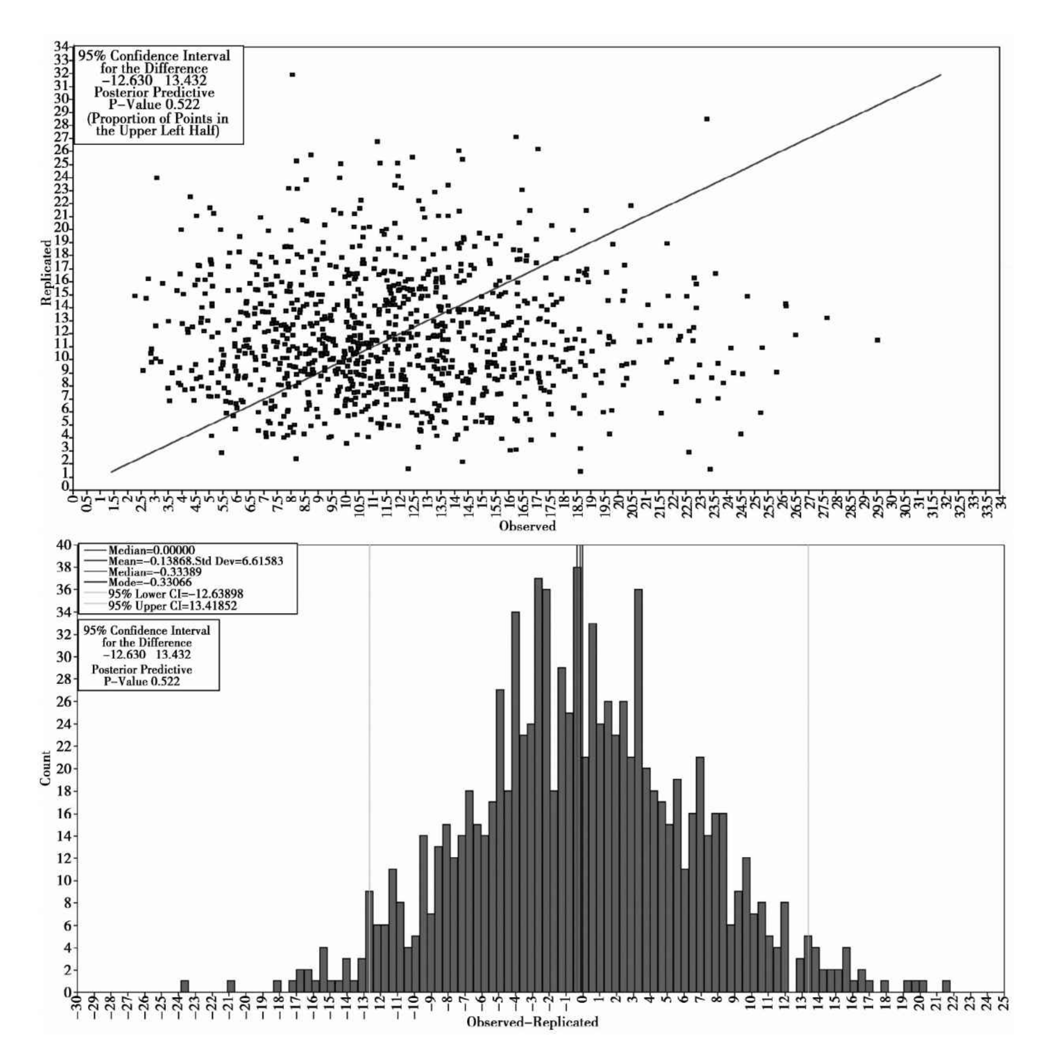


图10-11　中介效应分析的后验预测散点图和直方图

框10-5　中介分析结果（节选）

```
MODEL FIT INFORMATION

Number of Free Parameters                                   12

Bayesian Posterior Predictive Checking using Chi-Square
```

```
    95%Confidence Interval for the Difference Between
    the Observed and the Replicated Chi-Square Values
                    -12.630           13.432
    Posterior Predictive P-Value                    0.522
Information Criteria
    Deviance (DIC)                            57012.288
    Estimated Number of Parameters (pD)          11.754
    Bayesian (BIC)                            57087.637
STANDARDIZED MODEL RESULTS
STDYX Standardization
                      Posterior One-Tailed        95%C. I.
            Estimate       S. D.        P-Value  Lower   2.5%
Upper  2.5%  Significance
DEPRE      ON
  NEGTIVE           0.179       0.014       0.000       0.151
0.205      *
  STRESS            0.461       0.012       0.000       0.436
0.484      *
  SELF              0.003       0.014      0.418       -0.024
0.031
NEGTIVE    ON
  STRESS            0.204       0.016       0.000       0.173
0.234      *
SELF      ON
  STRESS           -0.175       0.016      0.000       -0.206
-0.143      *
NEGTIVE WITH
```

```
  SELF                  0.034       0.016       0.022       0.001
0.065      *
```

10.8.4　模型结果解释

这里我们并没有设定路径系数的先验分布，而且采用的样本量较大，近4000人，所以贝叶斯的结果和ML的结果差异不大，仅在小数点的第三位有细微差异，基本可以忽略。需要注意的是，贝叶斯估计对概率（可信区间）的解释是直接的，即使与ML估计的结果差异不大，但贝叶斯估计仍然有解释上的优势。如果在实践中，研究者对路径系数有适当的预期，可以设定有信息的先验。

10.9　贝叶斯混合模型

混合模型同时包含连续潜变量和类别潜变量，采用传统的估计方法通常非常耗时，而采用贝叶斯估计可以显著提高计算效率。

10.9.1　背景介绍和模型设定

在本书第4章 和第5章 介绍增长曲线模型和潜类别增长模型时均使用了中国健康与营养调查的数据，这里我们依然使用这个数据来演示采用贝叶斯估计GMM的分析过程。

关于GMM模型的设定，我们采用三类别自由估计时间分的模型（在实际应用中，模型的设定通常依据现有理论）。具体设定如下：

①线性增长，时间分自由估计且跨类别相等；

②增长截距和斜率因子的均值、方差和协方差跨类别自由估计；

③不同时间点测量的误差方差限定相等。

这里我们设定了增长截距和斜率因子均值和潜类别数目的先验信息。具体的先验信息设定如下[8]：

①增长截距和斜率因子均值正态分布，类别1～3的均值超参数分别为100，80和90，方差均设定为10。

②类别内计数服从Dirichlet分布，这里设定类别1和类别2的先验超参数分别为*D*（40，775）和*D*（50，775）[9]。

10.9.2　模型收敛检验

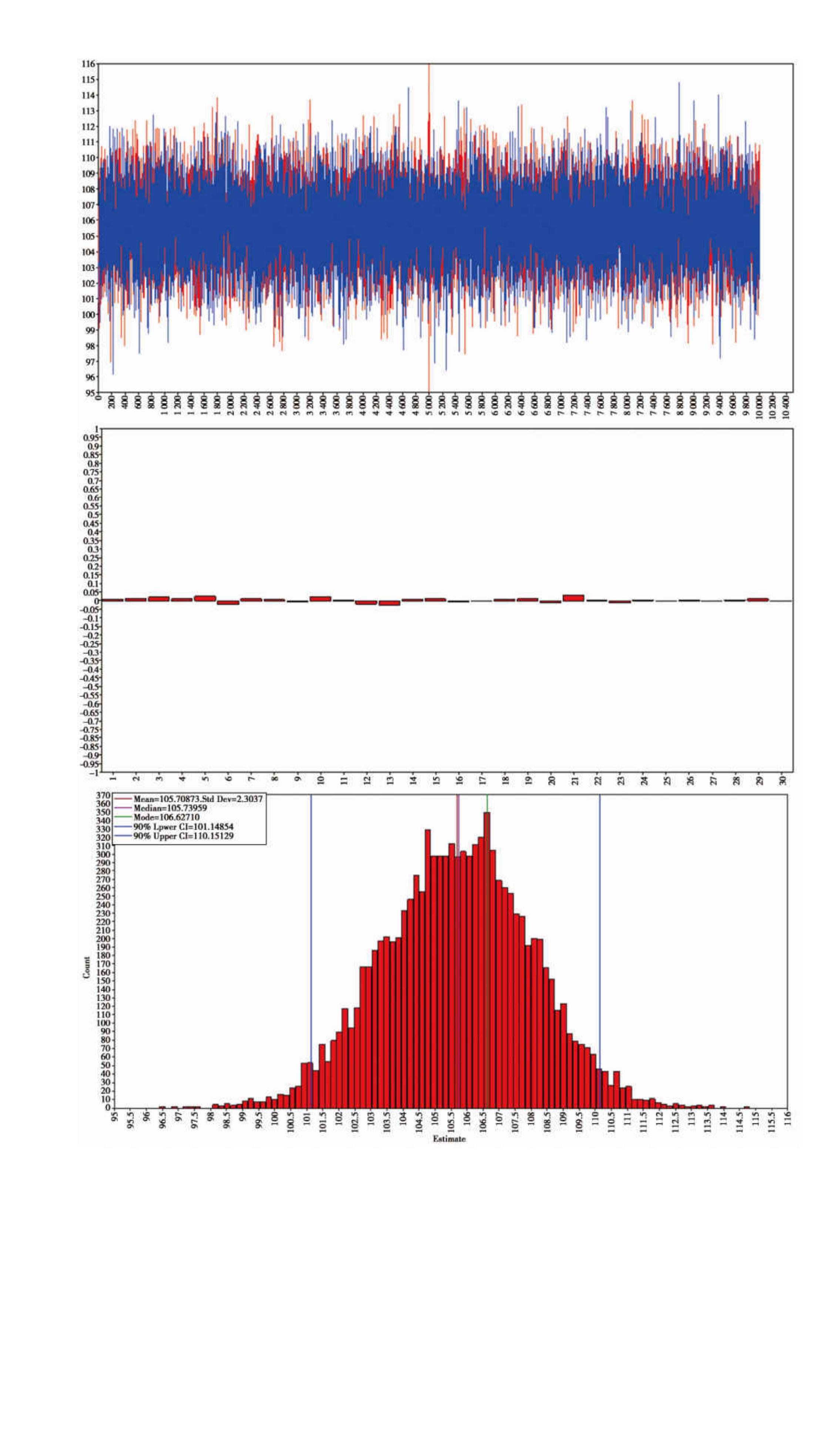
如果采用软件默认的PSR收敛标准，迭代7次即可收敛，但后验估计分布较差，因此采用固定迭代次数10000次，前5000次用作burn-in，后5000次用作后验估计。该模型的M*plus* 语句见框10-6。

框10-6　GMM的贝叶斯估计语句

```
Title: An example of GMMwith Bayes estimator;
Data: File is chnspress__wide. dat;
Variable: Names are id gender bmi89 bmi91 bmi93 bmi97 bmi00
bmi04 bmi06
         press89  press91  press93  press97  press00  press04
press06;
  Missing are all (－9999);
    USEVARIABLES ＝ press89  press91  press93  press97  press00
press04 press06;
    Classes＝c (3);
Analysis: Type＝Mixture;
      Estimator＝Bayes;
      FBITER＝10000; ! 固定迭代次数10000;
      Processor＝2;
      Chains＝2;
      point＝mean; ! 后验统计量估计的是均值，软件默认的是中位数;
      thin＝30;
Model:
      %Overall%
       i  s | press89@0  press91@1  press93@2  press97  press00
press04 press06;
     press89-press06 (1); ! 所有时间点测量的误差方差相等;
      i  s;
      [C#1*-1] (g);
      [C#2*0] (h);
      %C#1%
```

```
      [i] (a);
      [s] (b);
      i with s;
      i s;
      %C#2%
      [i] (c);
      [s] (d);
      i with s;
      i s;
      %C#3%
     [i] (e);
     [s] (f);
      i with s;
      i s;
  model priors: ! 设定模型参数的先验信息
    a~N(100, 10); ! 这些超参数是根据ML估计确定的，可以尝试增加或减
少这些
    b~N(-.5, 10); ! 数值做敏感性分析;
    c~N(80, 10);
    d~N(.5, 10);
    e~N(90, 10);
    f~N(.8, 10);
    g~D(40, 775);
    h~D(50, 775);
    Output: TECH8 stand;
      Plot: Type is Plot3;
        Series=press89-press06(*);
```

图10-12呈现了类别1的截距因子均值的踪迹图、链2的自相关系数图和参数估计后验分布图（其他参数结果与之类似）。踪迹图提示参数收敛到一个平稳的分布；自相关系数较小，说明抽样之间相互独立（注意这里使用了$thin$ =30）。参数估计的后验分布接近正态分布，所有这些结果提示模型参数收敛正常。



Mean=105.70873.Std Dev=2.3037
Median=105.73959
Mode=106.62710
90% Lpwer CI=101.14854
90% Upper CI=110.15129
Count
Estimate

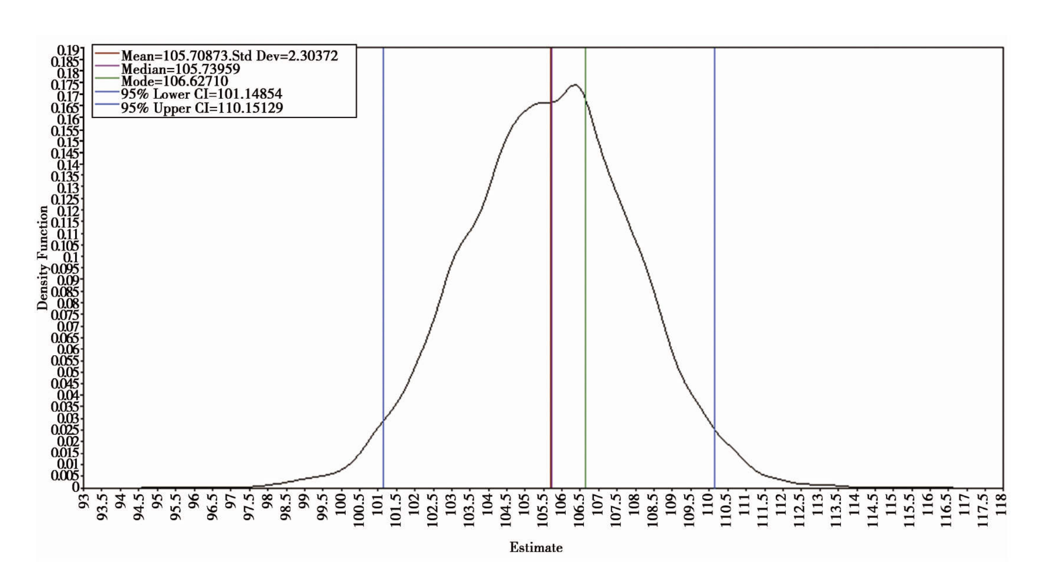


图10-12　类别1的截距因子均值的踪迹、自相关系数和参数估计后验分布图

10.9.3　模型拟合评价

图10-13呈现了模型的后验预测直方图和散点图。后验预测p 值等于0.000，说明模型拟合不理想，模型不能很好地复制观测数据。其中95％的可信区间为［103.352，211.235］，提示观测数据的卡方值大于后验分布生成数据的卡方值。散点图的结果进一步确认了这一判断，大部分散点落在45°线的下方，说明所有的p 值均较小，足以拒绝模型。

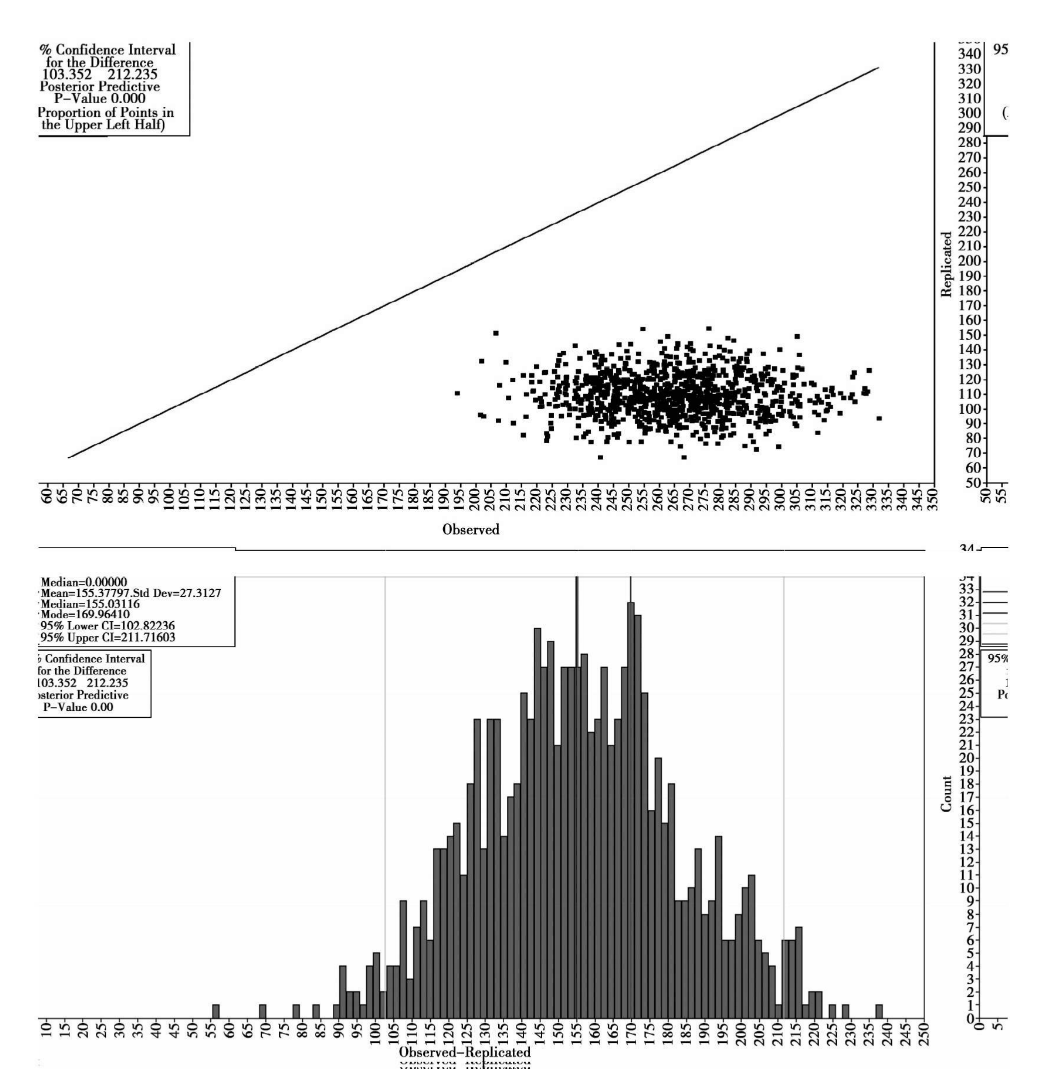


图10-13 GMM模型PPC预测直方图和散点图

10.9.4 模型结果解释

模型的部分结果整理在表10-5中。3个类别的比例分布为0.057，0.06和0.884，乘以总样本量即得到各类别的人数。类别1中的增长截距和斜率因子间的协方差为2.423，对应的95％的可信区间为［－0.755 8.416］。另外，该类别初始血压均值为105.709，平均增长率

为0.124；类别2和类别3的初始血压均值分别为85.646和87.451，平均增长率分别为－0.976和－0.234。

表10-5　GMM结果节选（非标准化结果）

	Estimate	S.D.	P-Value	95% C.I.	
Class1 Proportions	0.057	0.007	0.000	0.043	0.071
S WITH I	2.423	2.448	0.096	−0.755	8.416
I Means	105.709	2.304	0.000	101.149	110.151
S Means	0.124	0.147	0.168	−0.196	0.404
Class 2 Proportions	0.060	0.008	0.000	0.045	0.076
S WITH I	−1.361	1.142	0.058	−4.105	0.477
I Means	85.646	1.732	0.000	82.224	89.059
S Means	−0.976	0.305	0.000	−1.667	−0.471
Class 3 Proportions	0.884	0.010	0.000	0.863	0.903
S WITH I	−0.309	0.166	0.014	−0.682	−0.030
I Means	87.451	0.347	0.000	86.771	88.144
S Means	−0.234	0.070	0.000	−0.396	−0.115
Residual Variances[1]	81.842	2.068	0.000	77.905	85.918

[1] 设定了所有时间点和类别测量的误差方差相等。

类别1的个体起始血压水平较高同时随时间缓慢增加；后两个类别的起始血压水平较低，但随时间缓慢降低。

10.10　本章小结

本章系统（尽管算不上细致）介绍了贝叶斯分析的基本原理和过程，并采用几个潜变量建模的例子演示了分析的过程。但总的来说，同其他统计应用领域一样，贝叶斯在潜变量建模领域的应用还比较稀少，相关的方法学研究还需要深入。

尽管贝叶斯统计作为现代统计的一大流派，同频率学派平分天下，然而在应用统计领域的影响十分有限，在社科领域的应用更加稀

少。希望本章的介绍可以为感兴趣的读者提供一个新的数据分析思路和工具选择。

表10-6 M*plus* 常用贝叶斯分析语句

M*plus* syntax	备注
ANALYSIS：	
ESTIMATOR IS Bayes；	估计方法选择贝叶斯估计
POINT IS mean／median／mode；	选择贝叶斯的点估计：均值／中位数／众数。M*plus* 默认的是中位数。
BCONVERGENCE IS .05；	收敛标准，默认0.05，通常推荐0.01.
BITERATIONS IS a（b）.	与BCONVERGENCE配合使用，用以设定每条马尔科夫链最大（a）和最小（b）迭代次数；默认最大为50000，这里需要注意，当PSR小于1.05时迭代结束，如果想扩大迭代次数可以通过FBITERATIONS设定。
FBITERATIONS IS＃；	设定迭代次数。
CHAIN IS＃；	设定马尔科夫链的数目，默认2条。如果设置多个处理器PROCESSORS IS＃，每条链是独立运行的。
BSEED IS＃；	设定MCMC的起始点，很多软件默认的起始点是电脑的时钟；如果设定BSEED将会得到相同的结果。
STVALUES IS	为了提高计算时间可以将极大似然估计的结果作为

ml；	MCMC的起始值。
THIN IS＃；	“瘦身”“煲机”。由于迭代次数较多，如果每次迭代的结果都保留并计算后验分布将会非常耗时也会占用大量存储空间。特别是当前后两次迭代相关非常高时，此时可间隔一段距离，比如每隔10次。
MODEL：［C0score］（p1）；	给待估参数设定标签，目的是为其设定先验分布。
MODEL PRIORS：p1～N（＃a，＃b）；	设定先验分布。N表示正态分布，＃a和＃b分别表示均值和方差。
OUTPUT：STAND；CINTERCAL；	PPI区间。
PLOT：TYPE IS PLOT3；	报告轨迹图直方图和核密度图。

※推荐阅读※

Lynch的教科书写得非常好，通俗易懂，适合作为入门教材，当然对于大部分读者来说也已经足够。Gelman等的*Bayesian data analysis* 可以说是贝叶斯统计领域的圣经，非常细致，适合不同层次的读者。如果大家只想了解贝叶斯的大致情况以及在建模领域的应用，可以参考Muthén & Asparouhov（2012）和Zyphur & Oswald（2015）的导读性文章。

Gelman，A. ，Carlin，J. B. ，Stern，H. S. ，Dunson，D. B. ，Vehtari，A. ，＆Rubin，D. B. （2014）. Bayesian data analysis（3rd ed. ）. Boca Raton：CRC Press.

Lynch，S. M. （2010）. *Introduction to applied Bayesian statistics and estimation for social scientists* . New York：Springer.

Muthén，B. ，＆Asparouhov，T. （2012）. Bayesian structural equation modeling：A more flexible representation of substantive theory. *Psychological Methods* ，17，313-335.

Zyphur，M. J. ，＆Oswald，F. L. （2015）. Bayesian estimation and inference A User's Guide . *Journal of Management* ，41，390-420.

---

(1)　先验分布与似然值结合后得到的后验分布与先验分布同属一个分布家族，称作共轭先验（Conjugate Priors）。关于M*plus* 各种模型采用的先验分布可以参考（Asparouhov＆Muthén，2010），具体设定可在结果输出的TECH1中获得。另外，关于常用的先验分布可参见Lynch（2010）第3章。

(2)　贝叶斯统计里，由于参数的不确定性，在设置先验分布时需要确定分布的位置和离散信息。假设分布为正态分布，需要确定均值和方差，这些参数称为超参数（hyperparameters）。

(3)　这里不妨用生活中的例子来说明这一原理：天空乌云密布，电闪雷鸣，根据先前的生活经验我们判断100％会有一场暴雨，但是过了一会儿乌云散去，阳光再现；多日之后又出现类似的天气，我们判断下雨的概率不会是100％，而可能是90％，但这次同样也没有下雨；第三次出现类似情况时，我们判断下雨的概率会继续下降，比如80％。我们很多的主观概率都是来自生活经验的积累。

(4)　相同的迭代次数，M*plus* 的运算速度比BUGS更快（Wang＆Preacher，2015）。

(5) 在M*plus* 中，如果模型中所有参数的PSR小于1时，将会报告模型收敛了。

(6) 迭代次数都是50 000。

(7) 对应的模型参数是negative on stress完全标准化解。

(8) 这里的超参数设定均参考了“基础篇”第5章ML估计的结果。当然，这里主要用作展示的目的，如果在实际应用中，要考虑更可靠的信息来源。

(9) 如果是3个潜类别，则需要设定2个类别的先验参数，以此类推。通常，第一个数为该类别可能的计数，后一个数为总样本量数。

参考文献

Abroms，L. ，Simons-Morton，B. ，Haynie，D. L. ，＆Chen，R. （2005）. Psychosocial predictors of smoking trajectories during middle and high school. *Addiction，100，* 852-861.

Akremi，A. E. ，Colaianni，G. ，Portoghese，I. ，Galletta，M. ，＆Battistelli，A. （2014）. How organizational support impacts affective commitment and turnover among Italian nurses：a multilevel mediation model. *The International Journal of Human Resource Management，25，* 1185-1207.

Allison，P. D. （2009）. *Fixed Effects Regression Models*. Los Angeles：Sage.

Arbuckle，J. A. （1996）. *Full information estimation in the presence of incomplete data*. In G. A. Marcoulides＆R. E. Schumacker （Eds. ），Advanced structural equation modeling：Issues and techniques （pp. 243-277）. Hillsdale，NJ：Erlbaum.

Asparouhov，T. ，＆Muthén，B. （2003）. Full-information maximum-likelihood estimation of general two-level latent variable models. *Technical Report. Los Angeles*. Retrieved from http：//www.statmodel. com

Asparouhov，T. ，&Muthén，B. （2006）. Constructing covariates in multilevel regression. *Mplus Web Notes, 11* ，1-8. Retrieved from http：//www.statmodel. com.

Asparouhov，T. ，&Muthén，B. （2007）. Computationally efficient estimation of multilevel high-dimensional latent variable models. In *Proceedings of the 2007 JSM meeting in Salt Lake City, Utah, Section on Statistics in Epidemiology* （pp. 2531-2535）. The 2007 JSM meeting in Salt Lake City, Utah.

Asparouhov，T. ，&Muthén，B. （2008）. Multilevel mixture models. In G. R. Hancock&K. M. Samuelson （Eds. ），*Advances in latent variable mixture models* . Charlotte，NC：Information Age Publishing.

Asparouhov，T. ，&Muthén，B. （2010）. *Bayesian analysis using Mplus：Technical implementation* . Mplus Technical Report. Retrieved from http：//www.statmodel. com.

Asparouhov，T. ，&Muthén，B. （2012）. Multiple group multilevel analysis. *Mplus Web Notes＃16* . Retrieved from http：//www.statmodel. com.

Asparouhov T. &Muthén B. （2014）. Auxiliary variables in mixture modeling：Three-step approaches using Mplus. *Structural Equation Modeling, 21* ，329-341.

Bakk，Z. ，Oberski，D. L. ，＆Vermunt，J. K. （2016）. Relating latent class membership to continuous distal outcomes：improving the LTB approach and a modified three-step implementation. *Structural Equation Modeling, 23,* 278-289.

Bakk，Z. ，Tekle，F. B. ，＆Vermunt，J. K. （2013）. Estimating the association between latent class membership and external variables using bias adjusted three-step approaches. In T. F. Liao （Ed. ），*Sociological methodology* （pp. 272-311 ）. Thousand Oaks，CA：Sage.

Bakk，Z. ＆Vermunt，J. K. （2016）. Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling, 23,* 20-31.

Baldwin，S. A. ，＆Fellingham，G. W. （2013）. Bayesian methods for the analysis of small sample multilevel data with a complex variance structure. *Psychological Methods，18，* 151-164.

Barbour，J. B. ，＆Lammers，J. C. （2015 ）. Measuring professional identity：a review of the literature and a multilevel confirmatory factor analysis of professional，identity constructs. *Journal of Professions＆Organization，2,* 38-60.

Baron，R. M. ，&Kenny，D. A. （1987）. The moderator-mediator variable distinction in social psychological research : conceptual , strategic , and statistical considerations. *Journal of Personality&Social Psychology, 51,* 1173-1182.

Bartholomew，D. J. （1987）. *Latent variables models and factor analysis* . New York：Oxford University Press.

Bartholomew，D. J. ，&Knott，M. （1999）. *Latent variables models and factor analysis* （2nd ed. ）. London：Arnold.

Bauer，D. B. ，&Curran，P. J. （2004）. The integration of continuous and discrete latent variable models：Potential problems and promising opportunities. *Psychological Methods, 9,* 3-29.

Bauer，D. J. ，&Curran，P. J. （2003）. Distributional assumptions of growth mixture models：Implications for over-extraction of latent trajectory classes . *Psychological Methods, 8,* 338-363.

Bauer，D. J. （2003）. Estimating multilevel linear models as structural equation models. *Journal of Educational and Behavioral Statistics, 28,* 135-167.

Bauer，D. J. ，Preacher，K. J. ，&Gil，K. M. （2006）. Conceptualizing and testing random indirect effects and

moderated mediation in multilevel models: new procedures and recommendations. *Psychological Methods, 11,* 142-163.

Beauchamp, M. R. , Liu, Y. , Morton, K. L. , Martin, L. J. , Wilson, A. H. , &Wilson, A. J. , et al. (2014). Transformational teaching and adolescent physical activity: multilevel and mediational effects. *International Journal of Behavioral Medicine, 21,* 537-546.

Bentler, P. M. , &Liang, J. (2003). *Two-level mean and covariance structures: maximum likelihood via EM algorithm* , In S. P. Reise & N. Duan (Eds.), Multilevel Modeling: Methodological Advances, Issues, and Applications (pp. 53-70). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Bernstein, A. , Stickle, T. R. , Zvolensky, M. J. , Taylor, S. , Abramowitz, J. , & Stewart, S. (2010). Dimensional, Categorical, or Dimensional-Categories: Testing the Latent Structure of Anxiety Sensitivity Among Adults Using Factor-Mixture Modeling. *Behavior Therapy, 41,* 515-529.

Biesanz, J. C. , Deeb-Sossa, N. , Papadakis, A. A. , Bollen, K. A. , &Curran, P. J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods, 9,* 30-52.

Blakely，T. A. ，＆Woodward，A. J. （2000）. Ecological effects in multi-level studies. *Journal of Epidemiology & Community Health, 54,* 367-374.

Bliese，P. （2010）. Introductory and intermediate growth modeling. http：//www.statmodel.com / download / Topic％203-v.pdf.

Bliese，P. D. （2000）. Within-group agreement，non-independence，and reliability：Implications for data aggregation and analysis，In K. J. Klein＆S. W. Kozlowski（Eds.），*Multilevel theory，Research，and Methods in Organizations* （pp. 349-381）. San Francisco：Jossey-Bass.

Blozis，S. A. ，＆Cho，Y. I. （2008）. Coding and centering of time in latent curve models in the presence of inter individual time heterogeneity. *Structural Equation Modeling, 15,* 413-433.

Bolck，A. ，Croon，M. A. ，＆Hagenaars，J. A. （2004）. Estimating latent structure models with categorical variables：One-step versus three-step estimators. *Political Analysis, 12,* 3-27.

Bollen，K. A. ，＆Curran，P. J. （2006）. *Latent Curve Models：A Structural Equation Perspective*. Hoboken，NJ：Wiley.

Boulton，A．J．（2011）．*Fit Index Sensitivity in Multilevel Structural Equation Modeling*．PhD thesis，University of Kansas.

Bray，B. C. ，Lanza，S. T. ＆Tan，X. （2014）Eliminating Bias in Classify-Analyze Approaches for Latent Class Analysis. *Structural Equation Modeling，22*，1-11.

Bryk，A. S. ，＆Raudenbush，S. W. （1987）. Application of hierarchical linear models to assessing change . *Psychological Bulletin，101*，147-158.

Card N. A. （2012）. Multilevel mediational analysis in the study of daily lives. In M．R．Mehl，＆T．S．Conner（Eds．），*Handbook of Research Methods for Studying Daily Life* （pp. 479-494）. New York：Guilford.

Carragher，N. ，Adamson，G. ，Bunting，B. ，＆McCann，S. （2009）. Subtypes of depression in a nationally representative sample. *Journal ofAffective Disorders，113，* 88-99.

Chan，T．W．，＆Goldthorpe，J．H．（2007）．Social stratification and cultural consumption：Music in England. *European sociological review，23，* 1-19.

Chaux，E. ，Molano，A. ，＆Podlesky，P. （2009）. Socio-economic，socio-political and socio-emotional variables

explaining school bullying: a country-wide multilevel analysis. *Aggressive Behavior, 47,* 520-529.

Cheah, B. C. (2009). *Clustering Standard Errors or Modeling Multilevel Data?* Research Note. New York, NY: Columbia University.

Cheung, M. W. -L., Leung, K., & Au, K. (2006). Evaluating multilevel models in cross-cultural research: an illustration with social axioms. *Journal of Cross-Cultural Psychology, 37,* 522-541.

Cheung, M. W. -L. & Au, K. (2005). Applications of multilevel structural equation modeling to cross-cultural research. Structural Equation Modeling, 12, 598-619.

Christian, G., Jacob, B., Ginger, L., Saul, S., & Grenard, J. L. (2013). Analyzing latent state-trait and multiple-indicator latent growth curve models as multilevel structural equation models. *Frontiers in Psychology, 4,* 975.

Clark, S. L. Mixture Modeling with Behavioral Data. University of California, 2010.

Clark, S. L., & Muthén, B. (2009). *Relating latent class analysis results to variables not included in the analysis*. Retrieved from http://statmodel2.com/download/relatinglca.pdf.

Cohen，J. （1977）. Statistical power analysis for the behavioral sciences. *Journal of the American Statistical Association, 84,* 19-74.

Cohen，J. （1994）. The Earth Is Round （p＜. 05）. *American Psychologist, 49*, 997-1003.

Collins，L. M. ，＆Lanza，S. T. （2010）. *Latent class and latent transition analysis：With applications in the social, behavioral, and health sciences*. New York：Wiley.

Cronbach，L. J. （1976）. *Research on Classrooms and Schools：Formulation of Questions，Design and Analysis*. Stanford, CA：Stanford University Evaluation Consortium.

Cronbach，L. J. （1976）. *Research on classrooms and schools：Formulation of questions，design and analysis*. Occasional paper of the Stanford Evaluation Consortium，Stanford University，Stanford，CA.

Croon，M. A. ＆van Veldhoven，M. J. （2007）. Predicting group-level outcome variables from variables measured at the individual level：a latent variable multilevel model. *Psychological Methods, 12,* 45-57.

Curran，P. J. （2003）. Have multilevel models been structural equation models all along? *Multivariate Behavioral Research, 38*, 529-569.

Curran，P. J. ，＆Hussong，A. M. （2003）. The Use of latent Trajectory Models in Psychopathology Research. *Journal of Abnormal Psychology，112，* 526-544.

Davidov，E. ，Dülmer，H. ，Cieciuch，J. ，Kuntz，A. ，Seddig，D. ，＆Schmidt，P. （2016 ）. Explaining Measurement Nonequivalence Using Multilevel Structural Equation Modeling：The Case of Attitudes toward Citizenship Rights. *Sociological Methods＆Research* . 3-42.

Davidov，E. ，Dülmer，H. ，Schlüter，E. ，Schmidt，P. ，＆Meuleman， B . （2013）. Using a multilevel structural equation modeling approach to explain cross-cultural measurement noninvariance . *Journal of Cross-Cultural Psychology，43，* 558-575.

De，B. P. ，Wilson，M. ，＆Acton，G. S. （2005）. A conceptual and psychometric framework for distinguising categories and dimensions. *Psychological Review，112，* 129-158.

Depaoli，S. ，＆Clifton，J. P. （2015）. A Bayesian approach to multilevel structural equation modeling with continuous and dichotomous outcomes . *Structural Equation Modeling，22，* 1-25.

Depaoli，S. ，＆van de Schoot，R. （in press）. Improving transparency and replication in Bayesian statistics： the

wambs-checklist. *Psychological Methods* .

Dias，J. G. （2007）. Model Selection Criteria for Model-Based Clustering of Categorical Time Series Data：A Monte Carlo Study. In：Decker R. ，Lenz H. J. （Eds. ）*Advances in Data Analysis. Studies in Classification，Data Analysis，and Knowledge Organization* . Springer，Berlin，Heidelberg.

Dienes ， Z . （ 2011 ） . Bayesian versus orthodox statistics ： Which side are you on ? *Perspectives on Psychological Science* ，6，274-290.

Dixon，M. A. ，&Cunningham，G. B. （2006）. Data aggregation in multilevel analysis：a review of conceptual and statistical issues. *Measurement in Physical Education & Exercise Science，10，* 85-107.

Duncan，T. E. ，Duncan，S. C. ，&Strycker，L. A. （2006）. *An Introduction to Latent Variable Growth Curve Modeling：Concepts，Issues，and Applications* （2nd ed. ）. Mahwah NJ：Lawrence Erlbaum.

Dunn，E. C. ，Masyn，K. E. ，Jones，S. M. ，Subramanian，S. V. ，&Koenen，K. C. （2015 ）. Measuring psychosocial environments using individual responses：an application of multilevel factor analysis to examining students in schools. *Prevention Science* ，16，718-733.

Dyer, N. G. , Hanges, P. J. , and Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *The Leadership Quarterly, 16,* 149-167.

Eicker, F. (1967). Limit theorems for regressions with unequal and dependent errors. In K. Kojima (Ed.), *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability* (pp. 59-82). University of California Press.

Eid, M. , Langeheine, R. , &Diener, E. D. (2003). Comparing typological structures across cultures by Multigroup Latent Class analysis: A primer. *Journal ofCross-Cultural Psychology, 34,* 195-210.

Edwards, J. R. , &Lambert, L. S. (2007). Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis. *Psychological Methods, 12,* 1-22.

Enders, C. K. and Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12,* 121-138.

方杰，温忠麟，张敏强，任皓.（2014）. 基于结构方程模型的多层中介效应分析. 心理科学进展，22，530-539.

方杰，张敏强，邱皓政．（2010）．基于阶层线性理论的多层级中介效应．心理科学进展，18，1329-1338.

Feskens，R．，Hox，J．J．（2011）．Multilevel structural equation modeling for cross-cultural research：Exploring resampling methods to overcome small sample size problems. In E．Davidov，P．Schmidt，＆J．Billiet （Eds．），*Cross-cultural Analysis: Methods and Applications* （pp. 341-356）．Routledge.

Finch，W．H．，＆Bronk，K．C．（2011）．Conducting confirmatory latent class analysis using Mplus．*Structural Equation Modeling, 18*，132-151.

Franzini，L．，Caughy，M．，Spears，W．，＆Fernandez Esquer，M．E．（2005）．Neighborhood economic conditions，social processes，and self-rated health in low-income neighborhoods in Texas：a multilevel latent variables model. *Social Science＆Medicine, 61,* 1135-1150.

Gajewski，B．J．，Boyle，D．K．，Miller，P．A．，Oberhelman，F．，＆Dunton，N．（2010）．A multilevel confirmatory factor analysis of the practice environment scale：a case study. *Nursing Research, 59,* 147-153.

Gamoran，A．（1992）．The variable effects of high school tracking. *American Sociological Review, 57,* 812-828.

Geldhof，G. J. ，Preacher，K. J. ，＆Zyphur，M. J. （2014）. Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods，19，* 72-91.

Gelman，A. ，Carlin，J. B. ，Stern，H. S. ，Dunson，D. B. ，Vehtari，A. ，＆Rubin，D. B. （2014）. *Bayesian data analysis* （3rd ed. ）. Boca Raton：CRC Press.

Gelman，A. ，Carlin，J. B. ，Stern，H. S. ，＆Rubin，D. B. （2004）. *Bayesian data analysis* （2nd ed. ）. London，UK：Chapman＆Hall.

Gelman，A. ，＆Hill，J. （2006）*Data Analysis Using Regression and Multilevel／Hierarchical Models* . Cambridge University Press.

Goldstein，H. （2003）. *Multilevel Statistical Models* . Wiley.

Goldstein，H. ，＆Mcdonald，R. P. （1988）. A general model for the analysis of multilevel data. *Psychometrika，53，* 455-467.

Graham，J. W. ，Taylor，B. J. ，Olchowski，A. E. ，＆Cumsille，P. E. （2006）. Planned missing data designs in psychological research. *Psychological Methods，11，* 323-343.

Grant，A. M. ，&Berry，J. W. （2011）. The necessity of others is the mother of invention：Intrinsic and prosocial motivations，perspective taking，and creativity. *Academy of Management Journal，54，* 73-96.

Grimm，K. J. ，&Ram，N. （2009）. Nonlinear growth models in Mplus and SAS. *Structural Equation Modeling，16*，676-701.

Guo，G. &John H. （2004）. Longitudinal analysis for continuous outcomes：random effects models and latent trajectory models. In Melissa Hardy and Alan Bryman （eds.），*Handbook of Data Analysis* （pp. 347-369）. Sage Publications：London.

Hausman，J. A. （1976）. Specification tests in econometrics. *Econometrica：Journal of the Econometric Society，46，* 1251-1271.

Hedeker，D. &Gibbons，R. （2006）. *Longitudinal Data Analysis*. Hoboken，Hew Jersey：Wiley&Sons.

Henry，K. L. &Muthén，B. （2010）. Multilevel Latent Class Analysis：An application of Adolescent smoking typologies with individual and contextual predictors. *Structural Equation Modeling，17，* 193-215.

Hipp，J. R. ，&Bauer，D. J. （2006）. Local solutions in the estimation of growth mixture models. *Psychological*

Methods, 11, 36-53.

Hofmann, D. A. & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management, 24,* 623-641.

Holtmann, J., Koch, T., Lochner, K., & Eid, M. (2016). A comparison of ml, wlsmv, and bayesian methods for multilevel structural equation models in small samples: a simulation study. *Multivariate Behavioral Research, 51,* 1-20.

Hox, J. (2010). *Multilevel Analysis: Techniques and Applications*. Routledge Academic.

Hox, J., Maas, C. J., & Brinkhuis, M. J. (2010). The effect of estimation method and sample size in multilevel structural equation modeling. *Statistica Neerlandica, 64,* 157-170.

Hsu, H. Y., Kwok, O. M., Lin, J. H., & Acosta, S. (2015). Detecting misspecified multilevel structural equation models with common fit indices: a Monte Carlo study. *Multivariate Behavioral Research, 50,* 197-215.

Hsu, H. Y. (2009). *Testing the effectiveness of various commonly used fit indices for detecting misspecifications in*

multilevel structure equation models . Dissertations＆Theses-Gradworks.

Hsu，H. Y. ，Lin，J. H. ，Kwok，O. M. ，Acosta，S. ，＆ Willson，V. （2016）. The Impact of Intraclass Correlation on the Effectiveness of Level-Specific Fit Indices in Multilevel Structural Equation Modeling A Monte Carlo Study. *Educational and Psychological Measurement* .

Huber，P. J. （1967）. The behavior of maximum likelihood estimates under nonstandard conditions . In K . Kojima （Ed．），*Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability* （pp．221-233）. University of California Press.

Huhtala，M. ，Tolvanen，A. ，Mauno，S. ，＆Feldt，T. （2015 ）. The associations between ethical organizational culture，burnout，and engagement：a multilevel study. *Journal of Business and Psychology，30，* 399-414.

Hunter，A．M．，Muthén，B．O．，Cook，I．A．，＆ Leuchter，A．F．（2010）．Antidepressant response trajectories and quantitative electroencephalography （QEEG） biomarkers in major depressive disorder . *Journal of Psychiatric Research，44，* 90-98.

Huttenlocher，J. ，Haight，W. ，Bryk，A. ，Seltzer，M. ，＆Lyons，T. （1991）. Early vocabulary growth：relation to

language input and gender. *Developmental Psychology, 27,* 236-248.

Hwang, J., & Hopkins, K. (2012). Organizational inclusion, commitment, and turnover among child welfare workers: A multilevel mediation analysis. *Administration in Social Work, 36,* 23-39.

Jackman, S. (2004). Bayesian analysis for political research. *Annual Review ofPolitical Science, 7,* 483-505.

Johnson, M. K. (2002). Social origins, adolescent experiences, and work value trajectories during the transition to adulthood. *Social forces, 80*, 1307-1340.

Julian, Marc W. (2001). The consequences of ignoring multilevel data structures in nonhierarchical covariance modeling. *Structural Equation Modeling A Multidisciplinary Journal, 8*, 325-352.

Kaplan, D. (2000). *Structural equation modeling: Foundations and extensions*. Newbury Park, CA: Sage Publications.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90,* 773-795.

Kenny，D．A．，Korchmaros，J．D．，& Bolger，N．（2003）．Lower level mediation in multilevel models．*Psychological methods, 8,* 115-128.

Kim, S-Y. （2012）. Sample Size Requirements in Single- and Multiphase Growth Mixture Models：A Monte Carlo Simulation Study. *Structural Equation Modeling, 19,* 457-476.

Kim, E. S., Dedrick, R. F., Cao, C., &Ferron, J. M. （2016）. Multilevel factor analysis：reporting guidelines and a review of reporting practices. *Multivariate Behavioral Research, 51,* 881-898.

Klein, K. J., &Kozlowski, S. W. J. （2000）. From Micro to Meso：critical steps in conceptualizing and conducting multilevel research. *Organizational Research Methods, 3,* 211-236.

Kline, R. B. （2011）. Convergence of structural equation modeling and multilevel modeling. In M., Williams, &W. P. Vogt, （2011）. *The SAGE handbook of innovation in social research methods* （pp. 562-589）. Sage Publications.

Koss，K．J．，George，M．R．W．，Davies，P．T．，Cicchetti，D．，Cummings，E．M．，& Sturge-Apple，M．L．（2013）. Patterns of children's adrenocortical reactivity to interparental conflict and associations with child

adjustment: A growth mixture modeling approach. *Developmental Psychology, 49,* 317-326.

Kozlowski, S. W. & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel Theory, Research and Methods in Organizations: Foundations, Extensions, and New Directions* (pp. 3-90). San Francisco, CA: Jossey-Bass.

Kreft, I. G. (1997). The interactive effect of alcohol prevention programs in high school classes: An illustration of item homogeneity scaling and multilevel analysis techniques. In K. J. Bryant, M. Windle, and S. G. West (eds.), *Science of Prevention: Methodological Advances from Alcohol and Substance Abuse Research*. (pp. 251-278). Washington, DC: American Psychological Association.

Kreft, I. G. & De Leeuw, J. (1998). *Introducing Multilevel Modeling*. SAGE Publications Limited.

Kreft, I. G., De Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research, 30*, 1-21.

Kreft & Leeuwu (2007). 多层次模型分析导论. 邱皓政译. 重庆: 重庆大学出版社.

Krull，J. L. ＆MacKinnon，D. P. （1999）. Multilevel mediation modeling in group-based intervention studies. *Evaluation Review, 23,* 418-444.

Krull，J. L. ＆MacKinnon，D. P. （2001）. Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research, 36,* 249-277.

Kruschke，J. K. ，Aguinis，H. ，＆Joo，H. （2012）. The time has come Bayesian methods for data analysis in the organizational sciences. *Organizational Research Methods, 15,* 722-752.

Laird，N. M. ，＆Ware，J. H. （1982）. Random effects models for longitudinal data. *Biometrics, 38,* 963-974.

Lanza，S. T. ，Tan，X. ，＆Bray，B. C. （2013）. Latent class analysis with distal outcomes：A flexible model-based approach. S*tructural Equation Modeling, 20*，1-26.

Lanza，S. T. ＆Rhoades，B. L. （2013）. Latent Class Analysis：An alternative perspective on subgroup analysis in prevention and treatment. *Prevention Science, 14,* 157-168.

Lazarsfeld，P. F. ，＆Henry，N. W. （1968）. *Latent structure analysis*. Boston，MA：Houghton Mifflin.

Ledgerwood，A. ，＆Shrout，P. E. （2011）. The trade-off between accuracy and precision in latent variable models of

mediation processes . *Journal ofPersonality & Social Psychology，101，* 1174-1188.

Lee，V. E. ，＆Bryk，A. S. （1989）. A multilevel model of the social distribution of high school achievement . *Sociology of Education，62，* 172-192.

Liang，J. ，＆Bentler，P. M. （2004）. An EM algorithm for fitting two-level structural equation models . *Psychometrika，69，* 101-122.

廖卉，庄瑗嘉，陈晓萍，徐淑英，樊景立. （2012）. 多层次理论模型的建立及研究方法. 见：陈晓萍，徐淑英，樊景立（主编）. 组织与管理研究的实证方法（pp. 442-476）. 北京：北京大学出版社.

Li，L. ，＆Hser，Yih-Ing. （2011）. On inclusion of covariates for class enumeration of growth mixture models. *Multivariate Behavioral Research，46，* 266-302.

Li，X. ，＆Beretvas，S. N. （2013）. Sample size limits for estimating upper level mediation models using multilevel SEM. *Structural Equation Modeling，20*，241-264.

刘东，张震，汪默. （2012）. 被调节的中介和被中介的调节：理论构建与模型检验. 见：陈晓萍，徐淑英，樊景立（主编），组织与管理研究的实证方法（pp. 553-587）. 北京：北京大学出版社.

Lockwood，C. M.，& MacKinnon，D. P. （1998）. *Bootstrapping the standard error of the mediated effect*. In Proceedings of the 23rd annual meeting of SAS Users Group International （pp. 997-1002）.

Lo，Y.，Mendell，N.，& Rubin，D. B. （2001）. Testing the number of components in a normal mixture. *Biometrika, 88,* 767-778.

Lubke，G. H.，& Muthén，B. （2005）. Investigating population heterogeneity with factor mixture models. *Psychological Methods, 10,* 21-39.

Lubke，G. H. & Muthén，B. O. （2007）. Performance of factor mixture models as a function of covariate effects，model size，and class-specific parameters. *Structural Equation Modeling, 14,* 26-47.

Lubke，G. H.，& Neale，M. C. （2006）. Distinguishing between latent classes and continuous factors: Resolution by maximum likelihood? *Multivariate Behavioral Research，41*，499-532.

Lüdtke，O.，Marsh，H. W.，Robitzsch，A.，& Trautwein，U. （2011）. A 2 ×2 taxonomy of multilevel latent contextual models: accuracy-bias trade-offs in full and partial error correction models. *Psychological Methods，16,* 444-467.

Lüdtke，O. ，Marsh，H. W. ，Robitzsch，A. ，Trautwein，U. ，Asparouhov，T. ，＆Muthén，B. （2008）. The multilevel latent covariate model：a new，more reliable approach to group-level effects in contextual studies . *Psychological methods，13，* 203-229.

Lynch，S. M. （2010）. *Introduction to applied Bayesian statistics and estimation for social scientists* . New York：Springer.

MacCallum，R. C. ，Edwards，M. C. ，＆Cai，L. （2012）. Hopes and cautions in implementing Bayesian structural equation modeling. *Psychological Methods，17，* 340-345.

MacKinnon，D. P. ，Fritz，M. S. ，Williams，J. ，＆Lockwood，C. M. （2007）. Distribution of the product confidence limits for the indirect effect：Program PRODCLIN. *Behavior Research Methods，39* ，384-389.

Magidson，J. ，＆Vermunt，J. K. （2001 ）. Latent class factor and cluster models，biplots and related graphical displays. *Sociological Methodology，31* ，223-264.

Magidson，J. ，＆Vermunt，J. K. （2002）. Latent class models for clustering：A comparison with K-means. *Canadian Journal ofMarketing Research，20，* 37-44.

Marsh，H. W. ，Lüdtke，O. ，Robitzsch，A. ，Trautwein，U. ，Asparouhov，T. ，Muthén，B. ，＆Nagengast，B. （2009）.

Doubly-latent models of school contextual effects : Integrating multilevel and structural equation approaches to control measurement and sampling error . *Multivariate Behavioral Research, 44,* 764-802.

Marsh Herbert W. , Oliver Lüdtke, Benjamin Nagengast, Ulrich Trautwein , Alexandre J . S . Morin , & Adel S . Abduljabbar , et al . (2012) . Classroom climate and contextual effects: conceptual and methodological issues in the evaluation of group-level effects . *Educational Psychologist* , 47 (2) , 106-124.

Masyn, K. , Henderson, C. , &Greenbaum, P. (2010) . Exploring the latent structures of psychological constructs in social development using the Dimensional-Categorical Spectrum. *Social Development, 19,* 470-493.

McArdle, J. J. , &Bell, R. Q. (2000) . An introduction to latent growth models for developmental data analysis. In Little, T. D. , Schnabel, K. U. , &Baumert, J. (Eds) . *Modeling Longitudinal and Multiple-Group Data : Practical issues, applied approaches, and scientific examples* . (pp. 69-107) . Mahwah, NJ: Erlbaum.

McCutcheon, A. L. (2002) . Basic concepts and procedures in single-and multiple-group latent class analysis. In J. A. Hagenaars&A. L. McCutcheon (Eds.) . *Applied latent class*

analysis （pp. 57-88 ）. Cambridge，UK：Cambridge University Press.

McDonald，R. P. （1994）. The bilevel reticular action model for path analysis with latent variables. *Sociological Methods and Research，22，* 399-413.

McGrath，R. E. （2008）. Inferential errors in taxometric analyses of ordered three-class constructs . *Journal of Personality Assessment，90，* 11-25.

McLachlan，G. J. ，＆Peel，D. （2000）. *Finite mixture models* . New York：Wiley.

McNeish （ 2017 ） . Multilevel Mediation with Small Samples ： A Cautionary Note on the Multilevel Structural Equation Modeling. *Structural Equation Modeling* .

Meehl，P. E. ，＆Yonce，L. J. （1994）. Taxometric analysis： Ⅰ. Detecting taxonicity with two quantitative indicators using means above and below a sliding cut （MAMBAC procedure）. *Psychological Reports，74* ，1059-1274.

Meehl，P. E. ，＆Yonce，L. J. （1996）. Taxometric analysis： Ⅱ. Detecting taxonicity using covariance of two quantitative indicators in successive intervals of a third indicator （MAXCOV procedure）. *Psychological Reports，78* ，1091-1227.

Mehta，P．D．＆Neale，M．C．（2005）．People are variables too：multilevel structural equations modeling．*Psychological Methods，10，* 259-284.

Meng，X. L. （1994）Posterior predictive p-values. *The Annals ofStatistics，22，* 1142-1160.

Meredith，W．，＆Tisak，J．（1990）．Latent curve analysis. *Psychometrika，55，* 107-122.

Morin，Alexandre J．S．，Herbert W．Marsh，Benjamin Nagengast，＆L．Francesca Scalas．（2014）．Doubly latent multilevel analyses of classroom climate：an illustration. *Journal of Experimental Education，82（2）*，143-167.

Morin，A. J.，Morizot，J.，Boudrias，J. S.，＆Madore，I．（2011）．A multifoci person-centered perspective on workplace affective commitment：A latent profile / factor mixture analysis. *Organizational Research Methods，14，* 58-90.

Morris，N. A.，＆Slocum，L. A. （2012）．Estimating Country-Level Terrorism Trends Using Group-Based Trajectory Analyses：Latent Class Growth Analysis and General Mixture Modeling. *Journal ofQuantitative Criminology，28，* 103-139.

Muthén，B. （1984）. A general structural equation model with dichotomous，ordered categorical，and continuous latent variable indicators. *Psychometrika，49，* 115-132.

Muthén，B.，& Asparouhov，T.（2011）. Beyond multilevel regression modeling：Multilevel analysis in a general latent variable framework. In J. J. Hox & J. K. Roberts（Eds.），*Handbook for advanced multilevel analysis*.（pp. 15-40）. New York，NY US：Routledge / Taylor & Francis Group.

Muthén，B.，& Asparouhov，T.（2012）. Bayesian structural equation modeling：A more flexible representation of substantive theory. *Psychological Methods，17，* 313-335.

Muthén，B.（2003）. Statistical and substantive checking in growth mixture modeling. *Psychological Methods，8，* 369-377.

Muthén，B.（2006）. Should substance use disorders be considered as categorical or dimensional？*Addiction，101*，6-16.

Muthén，B.（2008）. Latent variable hybrids. In G. R. Hancock，& K. M. Samuelson（Eds.），*Advances in latent variable mixture models*（pp. 1-24）. Charlotte，NC：Information Age Publishing.

Muthén，B.（2010）. *Bayesian analysis in Mplus：A brief introduction*. Mplus Technical Report. Retrieved from http：//www. statmodel.com.

Muthén，B. ＆Brown，H. （2009）. Estimating drug effects in the presence of placebo response：Causal inference using growth mixture modeling. *Statistics in Medicine，28，* 3363-3385.

Muthén，B. ，＆Muthén，L. （2000）. Integrating person-centered and variable-centered analyses：Growth mixture modeling with latent trajectory classes. *Alcoholism：Clinical and Experimental Research，24，* 882-891.

Muthén L. K，＆Muthén B. O. （2012）. *Mplus User's Guide* . Los Angeles，CA：Muthén＆Muthén.

Muthén，B. O. （1989）. Latent variable modeling in heterogeneous populations. *Psychometrika，54，* 557-585.

Muthén，B. O. （1990）. *Mean and covariance structure analysis ofhierarchical data* . UCLA Statistics Series，62.

Muthén，B. O. （1991 ）. Multilevel factor analysis of class and student achievement components . *Journal of Educational Measurement，28，* 338-354.

Muthén，B. O. （1994）. Multilevel covariance structure analysis. *Sociological Methods＆Research，22，* 376-398.

Muthén，B. O. （2001）. Second-generation structural equation modeling with a combination of categorical and continuous latent variables：New opportunities for latent

class / latent growth modeling. In A. Sayer & L. Collins (Eds.), *New methods for the analysis of change* (pp. 291-322). Washington, DC: American Psychological Association.

Muthén, B. O., & Asparouhov, T. (2008). Growth mixture modeling: Analysis with non-Gaussian random effects. In G. Fitzmaurice, M. Davidian, G. Verbeke, & G. Molenberghs (Eds.), *Advances in longitudinal data analysis* (pp. 143-165). Boca Raton, FL: Chapman & Hall / CRC Press.

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology, 25,* 267-316.

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EMalgorithm. *Biometrics, 55,* 463-469.

Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group based approach. *Psychological Methods, 4,* 139-157.

Neale, M. C. (2000). *Individual fit, heterogeneity, and missing data in multigroup structural equation modeling*. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling Longitudinal and Multilevel Data: Practical Issues, Applied Approaches and Specific Examples* (pp. 249-267). Hillsdale, NJ: Erlbaum.

Neuhaus，J. M. &McCulloch，C. E. （2006）. Separating between-and within-cluster covariate effects by using conditional and partitioning methods. *Journal of the Royal Statistical Society: Series B （Statistical Methodology），68,* 859-872.

Nohe，C. ，Michaelis，B. ，Menges，J. I. ，Zhang，Z. ，& Sonntag，K. （2013）. Charisma and organizational change：A multilevel study of perceived charisma，commitment to change，and team performance. *The Leadership Quarterly，24,* 378-389.

Nylund，K. L. ，Asparouhov，T. ，&Muthén，B. O. （2007）. Deciding on the number of classes in latent class analysis and growth mixture modeling：A Monte Carlo simulation study. *Structural Equation Modeling，14，* 535-569.

O'Brien，R. M. （1990）. Estimating the reliability of aggregate-level variables based on individual-level characteristics. *Sociological Methods & Research，18*，473-504.

Oshio，T. （2016）. Which is more relevant for perceived happiness，individual-level or area-level social capital？A multilevel mediation analysis. *Journal of Happiness Studies*，1-19.

Paccagnella，O. （2006）. Centering or not centering in multilevel models？The role of the group mean and the assessment of group effects. *Evaluation Review，30，* 66-85.

Paulin，D.，&Griffin，B. （2016）. The relationships between incivility，team climate for incivility and job-related employee well-being：a multilevel analysis. *Work & Stress，30，* 1-20.

Petras，H.，&Masyn，K. （2010）. General growth mixture analysis with antecedents and consequences of change. In R. Piquero & D. Weisburd （Eds.），*Handbook of quantitative criminology* （pp. 69-100）. New York，NY：Springer.

Peugh，J. & Fan，X. （2012）. How well does growth mixture modeling identify heterogeneous growth trajectories? A simulation study examining gmm's performance characteristics. *Structural Equation Modeling，19，* 204-226.

Pituch，K.，Murphy，D.，&Tate，R. （2009）. Three-level models for indirect effects in school and class-randomized experiments in education. *The Journal of Experimental Education，78，* 60-95.

Pituch，K.，&Stapleton，L. （2011）. Hierarchical linear and structural equation modeling approaches to mediation analysis in randomized field experiments. In Williams，M.，&Vogt，W. P. （Eds.），*The SAGE handbook of*

innovation in social research methods （pp. 590-619）. Sage Publications.

Preacher，K. J. ＆Selig，J. P. （2012）. Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures，6，* 77-98.

Preacher，K. J. ，Zhang，Z. ，＆Zyphur，M. J. （2011）. Alternative methods for assessing mediation in multilevel data：The advantages of multilevel sem. *Structural Equation Modeling，18，* 161-182.

Preacher，K. J. ，Zhang，Z. ，＆Zyphur，M. J. （2016）. Multilevel structural equation models for assessing moderation within and across levels of analysis. *Psychological Methods，21，* 189-205.

Preacher，K. J. ，Zyphur，M. J. ，＆Zhang，Z. （2010）. A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods，15，* 209-233.

Preacher，K. J. （2011 ）. Multilevel SEM strategies for evaluating mediation in three-level data. *Multivariate Behavioral Research，46，* 691-731.

Preacher，K. J. （2015）. Advances in mediation analysis：A survey and synthesis of new developments. *Annual Review of Psychology，66，* 825-852.

邱皓政.（2008）. *潜在类别模型的原理与技术* . 北京：教育科学出版社.

Qureshi，I. ，&Fang. （2011 ）. Socialization in open source software projects ： A growth mixture modeling approach. *Organizational Research Methods，14* ，208-238.

Rabe-Hesketh，S. ，&Skrondal，A. （2012）. *Multilevel and longitudinal modeling using Stata* （3rd ed）. STATA press.

Rabe-Hesketh，S. ，Skrondal，A. ，&Pickles，A. （2004）. Generalized multilevel structural equation modeling. *Psychometrika，69* ，167-190.

Raftery，A. E. （1986）Choosing models for cross-classifications. *American Sociological Review，51，* 145-146.

Rasbash，J. ，Steele，F. ，Browne，W. J. and Goldstein，H. （2009）. *A User's Guide to MlwiN* . Centre for Multilevel Modelling，University of Bristol.

Ratanavaraha，V. ，Jomnonkwao，S. ，Khampirat，B. ，Watthanaklang，D. ，&Iamtrakul，P. （2016）. The complex relationship between school policy，service quality，satisfaction，and loyalty for educational tour bus services：a multilevel modeling approach. *Transport Policy，45，* 116-126.

Raudenbush＆Bryk （2007）．分层线性模型：应用与数据分析方法．郭志刚等译．北京：社会科学文献出版社．

Raudenbush，S．W．，Bryk，A．S．，＆Cheong，Y．F．（2008）．*HLM6.06 for Windows* ［computer software］．Lincolnwood，IL：Scientific Software International．

Raudenbush，S．W．＆Bryk，A．S．（2002）．*Hierarchical Linear Models：Applications and Data Analysis Methods*．SAGE Publications，Incorporated．

Raudenbush，S．W．，＆Chan，W．S．（1993）．Application of a hierarchical linear model to the study of adolescent deviance in an overlapping cohort design．*Journal ofConsulting＆Clinical Psychology，61，* 941-951．

Raudenbush，S．W．（2001）．Comparing personal trajectories and drawing causal inferences from longitudinal data．*Annual Review ofPsychology，50，* 501-525．

Raudenbush，S．W．，Rowan，B．，＆Kang，S．J．（1991）．A multilevel，multivariate model for studying school climate with estimation via the EM algorithm and application to U．S．high-school data．*Journal of Educational and Behavioral Statistics，16，* 295-330．

Reise，S．P．，Ventura，J．，Nuechterlein，K．H．，＆Kim，K．H．（2005）．An illustration of multilevel factor analysis．*Journal of Personality Assessment，84，* 126-136．

Rhemtulla，M. ，&Little，T. D. （2012）. Planned Missing Data Designs for Research in Cognitive Development. *Journal of Cognition and Development, 13,* 425-438.

Robinson，W. S. （1950）. Ecological correlations and the behavior of individuals. *American Sociological Review, 15,* 351-357.

Ross，C. E. ，Reynolds，J. R. ，&Geis，K. J. （2000）. The contingent meaning of neighborhood stability for residents' psychological well-being. *American Sociological Review, 65,* 581-597.

Ryu，E. （2008）. *Evaluation of model fit in multilevel structural equation modeling： level-specific model fit evaluation and the robustness to non-normality*. Dissertations＆Theses-Gradworks.

Ryu，E. （2014）. Model fit evaluation in multilevel structural equation models. *Frontiers in Psychology, 5,* 81.

Ryu，E. &West，S. G. （2009）. Level-specific evaluation of model fit in multilevel structural equation modeling. *Structural Equation Modeling, 16,* 583-601.

Sampson，R. J. ，Raudenbush，S. W. ，&Earls，F. （1997）. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science, 277,* 918-924.

Sampson，R. J. ，＆Sharkey，P. （2008）. Neighborhood selection and the social reproduction of concentrated racial inequality. *Demography, 45* , 1-29.

Satorra，A. （2000）. *Scaled and Adjusted Restricted Tests in Multi-sample Analysis of Moment Structures* . Springer.

Satorra，A. and Bentler，P. M. （2001 ）. A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66* , 507-514.

Satorra，A. and Muthén，B. （1995）. Complex sample data in structural equation modeling. *Sociological Methodology, 25,* 267-316.

Schermellehengel，K. ，Kerwer，M. ，＆Klein，A. G. （2014）. Evaluation of model fit in nonlinear multilevel structural equation modeling. *Frontiers in Psychology，5,* 181.

Schweig，J. D. （2014）. Multilevel factor analysis and student ratings of instructional practice. UCLA Electronic Theses and Dissertations.

Selig，J. P. ，Card，N. A. ，＆Little，T. D. （2008）. Latent variable structural equation modeling in cross-cultural research: Multigroup and multilevel approaches. In van de Vijver，Fons J. R. ，van Hemert，Dianne A. ＆

Poortinga，Ype H．（Eds．），*Multilevel Analysis of Individuals and Cultures*（pp. 93-119）. New York，NY：Taylor ＆Francis Group / Lawrence Erlbaum Associates

Singer，J. D.，＆Martin，N. C.（1998）. The design and analysis of longitudinal studies of development and psychopathology in context：statistical models and methodological recommendations. *Development and Psychopathology，10，* 395-426.

Singer，J. D. ＆Willett，J. B.（2003）. *Applied longitudinal data analysis：Modeling change and event occurrence*. Oxford university press.

Skrondal，A.，＆Rabe-Hesketh，S.（2004）. *Generalized latent variable modeling：Multilevel，Longitudinal，and Structural Equation Models*. CRC Press.

Snijders，T. A. ＆Bosker，R. J.（1999）. *Multilevel analysis：An introduction to basic and advanced multilevel modeling*. Sage Publications Limited.

Snijders，T. A. ＆Bosker，R. J.（2012）. *Multilevel analysis：An introduction to basic and advanced multilevel modeling*. Sage Publications Limited.

宋时歌，陈华珊.（2005）. 纵贯性数据与生长模型在社会科学实证研究中的应用. 社会学研究（5），69-91.

Spiegelhalter, D. J. , Best, N. G. , Carlin, B. P. , & van der Linde , A . (2002) . Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B (Statistical Methodology), 64*, 583-639.

Stapleton, L. M. (2013). Multilevel structural equation modeling with complex sample data. In Gregory R. Hancock & Ralph O. Mueller (Eds.), *Structural Equation Modeling: A Second Course* (pp. 521-562), Information Age Publishing Inc.

Stapleton, Laura M. , Yang, Ji Seung, Hancock, Gregory R . (2016) . Construct meaning in multilevel settings . *Journal of Educational&Behavioral Statistics, 41,* 481-520.

Steele, F. (2008). Multilevel models for longitudinal data . *Journal of the Royal Statistical Society : Series A (Statistics in Society), 171,* 5-19.

Sterba , S . K . (2013) . Understanding linkages among mixture models. *Multivariate Behavioral Research, 48,* 775-815.

Stoel, Reinoud D. , Godfried van Den Wittenboer & Joop Hox (2003) . Analyzing longitudinal data using multilevel regression and latent growth curve analysis. *Metodologia de las Ciencas del Comportamiento, 5,* 1-21.

Talloen，W. ，Moerkerke，B. ，Loeys，T. ，De Naeghel，J. ，Van Keer，H. ，&Vansteelandt，S. （2016）. Estimation of Indirect Effects in the Presence of Unmeasured Confounding for the Mediator-Outcome Relationship in a Multilevel 2-1-1 Mediation Model . *Journal of Educational and Behavioral Statistics，41，* 359-391.

Toit，S. H. C. D. ，&Toit，M. D. （2008）. Multilevel Structural Equation Modeling. In De Leeuw J. &Meijer E. （Eds. ），*Handbook ofMultilevel Analysis* （pp. 435-478），New York，NY：Springer.

Tofighi，D. （2010）. *Multilevel Mediation Analysis：Statistical Assumptions and Centering* . Doctoral dissertation，Arizona State University.

Tofighi，D. ，&Mackinnon，D. P. （2011）. RMediation：an R package for mediation analysis confidence intervals . *Behavior Research Methods，43，* 692-700.

Tofighi，D. ，&Kelley，K. （2016）. Assessing omitted confounder bias in multilevel mediation models. *Multivariate Behavioral Research，51，* 86-105.

Tofighi，D. ，West，S. G. ，&Mackinnon，D. P. （2013）. Multilevel mediation analysis：the effects of omitted variables in the 1-1-1 model. *British Journal ofMathematical and Statistical Psychology，66，* 290-307.

Tofighi，D. ，＆Thoemmes，F. （2014）. Single-level and multilevel mediation analysis . *The Journal of Early Adolescence，34，* 93-119.

Uthman，O. A. ，Moradi，T. ，＆Lawoko，S. （2010）. Are individual and community acceptance and witnessing of intimate partner violence related to its occurrence ? Multilevel structural equation model. *Plos One，6，* e27738.

Van de Schoot，R. ，Winter，S. ，Ryan，O. ，Zondervan-Zwijenburg，M. A. J. ，＆Depaoli，S. （in press）. A Systematic Review of Bayesian Papers in Psychology：The Last 25 Years. *Psychological Methods* .

Van Horn，M. L. ，Jaki，T. ，Masyn，K. ，Ramey，S. L. ，Smith，J. A. ，＆Antaramian，S. （2009）. Assessing differential effects：Applying regression mixture models to identify variations in the influence of family resources on academic achievement. *Developmental Psychology，45，* 1298-1313.

Van Kollenburg，Van G. H. （2011）. *Analysis of an Indirect （Mediated）Moderation Model* . Tilburg University Master Dissertation.

Van Kollenburg，G. H. ，＆Croon，M. A. （2017）. *Analysis ofan indirect （mediated） moderation model* . https：//arxiv.org/abs/1701.08862.

Vermunt，J. K. （2003）. Multilevel latent class models. *Sociological Methodology，33，* 213-239.

Vermunt，J. K. （2010）. Latent class modeling with covariates：Two improved three-step approaches. *Political Analysis，18，* 450-469.

Walsh，B. M.，&Matthews，R. A. （2010）. A multilevel model of the effects of equal opportunity climate on job satisfaction in the military. *Journal ofOccupational Health Psychology，15，* 191-207.

Walton，K. E.，Ormel，J.，&Krueger，R. F. （2011）. The Dimensional Nature of Externalizing Behaviors in Adolescence：Evidence from a Direct Comparison of Categorical，Dimensional，and Hybrid Models. *Journal of Abnormal Child Psychology，39，* 553-561.

Wang C. P.，Brown，C. H.，Bandeen-Roche，K. （2005）. Residual diagnostics for growth mixture models：Examining the impact of preventive intervention on multiple trajectories of aggressive behavior. *Journal of the American Statistical Association，100，* 1054-1076.

王济川，谢海义，姜宝法. （2008）. 多层统计分析模型：方法与应用. 北京：高等教育出版社.

Wang，L.，&Preacher，K. J. （2015）. Moderated mediation analysis using Bayesian methods. *Structural*

Equation Modeling，22， 249-263.

王孟成．（2014）．潜变量建模与Mplus应用：基础篇．重庆：重庆大学出版社．

王孟成，邓俏文，毕向阳，叶浩生，杨文．（2017）．潜在剖面模型的分类精确性：一项蒙特卡罗模拟研究．心理学报，49，1473-1482.

王孟成，毕向阳，叶浩生．（2014）．增长混合模型———分析不同类别个体发展趋势．社会学研究，4，220-241.

Wang，M．，＆Hanges，P．J．（2011）．Latent class procedures：Applications to organizational research．*Organizational Research Methods，14，* 24-31.

Wardell，J．D．，Ramchandani，V．A．，＆Hendershot，C．S．（2015）．A multilevel structural equation model of within-and between-person associations among subjective responses to alcohol，craving，and laboratory alcohol self-administration．*Journal ofAbnormal Psychology，124*，1050-1063.

温福星．（2009）．阶层线性模型的原理与应用．北京：中国轻工业出版社．

温忠麟，刘红云，侯杰泰．（2012）．调节效应和中介效应分析．北京：教育科学出版社．

White，H．（1980）．A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*，48，817-838.

伍德里奇.（2014）. 计量经济学导论. 费剑平译校. 北京：中国人民大学出版社.

Yang，C．（2006）. Evaluating latent class analyses in qualitative phenotype identification . *Computational Statistics＆Data Analysis，50，* 1090-1104.

Yuan，K．H．，＆Bentler，P．M．（2007）. Multilevel covariance structure analysis by fitting multiple single-level models. *Sociological Methodology，37*，53-82.

Yuan，Y．，＆MacKinnon，D．P．（2009）. Bayesian mediation analysis. *Psychological Methods，14，* 301-322.

Yuan，K．H．，＆Hayashi，K．（2005 ）. On Muthén's maximum likelihood for two－level covariance structure models. *Psychometrika*，70（1），147-167.

Yung，Y．F．（1997）. Finite mixtures in confirmatory factor analysis models. *Psychometrika，62，* 297-330.

张洁婷，焦璨，张敏强.（2010）. 潜在类别分析技术在心理学研究中的应用. 心理科学进展，18，1991-1998.

Zhang，Z．，Lee，C．K．，＆Ping，H．W．（2016）．Multilevel structural equation modeling analysis of the servant leadership construct and its relation to job satisfaction. *Leadership＆Organization Development Journal, 37,* 1147-1167.

Zhang，Z．，Zyphur，M．J．，and Preacher，K．J．（2009）．Testing multilevel mediation using hierarchical linear models problems and solutions. *Organizational Research Methods, 12,* 695-719.